DOCUMENT RESUME

ED 151 168 SE 023 801

AUTHOR	Ikeda, Hiroshi
TITLE	A Factorial Study of the Relationships Between Teacher-Held Objectives and Student Performance in UICSM High School Mathematics. University of Illinois Committee on School Mathematics (UICSM) Research Report No. 10.
INSTITUTION	Illinois Univ., Urbana. Committee on School Mathematics.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Jun 65
NOTE	203p.; For related document, see SE 023 800; Ph.D. Dissertation, University of Illinois; Not available in hard copy due to marginal legibility of original document
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Academic Achievement; Classroom Research; *Curriculum Evaluation; *Doctoral Theses; Educational Research; Mathematics Education; *Mathematics Instruction; Secondary Education; Secondary School Mathematics; *Statistical Analysis
IDENTIFIERS	*Research Reports

ABSTRACT

Two tests based on chapters 1-3 and 4-5, respectively, of "High School Mathematics, Course 1" by M. Beberman and H. E. Vaughan (1964) were given to 154 ninth-grade students at Pekin High School in Illinois before (pre-test) and after (post-test) coverage of chapters 1-3 and 4-5 in class. The tests were sent to 105 teachers from 70 schools in 19 states using this University of Illinois Committee on School Mathematics (UICSM) text. Teachers were asked to evaluate the suitability of each test question for their students. A factorial-analytic method was used to compute the intercorrelations among the pre-test scores, post-test scores, gain in test scores, and teacher rating for each of the 54 test questions. No particular relationships were found between mean values of the teacher ratings and the student performance on either pre-test items or post-test items, even when the mean values for Pekin teachers only were considered. A weakly positive relationship was found between mean values of the teacher ratings and the student gains. (Author/BB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151168

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

UICSM

RESEARCH REPORT

UNIVERSITY OF ILLINOIS COMMITTEE ON SCHOOL MATHEMATICS

2

SE 023 801

University of Illinois Committee on School Mathematics

Max Beberman, *Director*
1210 West Springfield
Urbana, Illinois

UICSM

RESEARCH REPORT

No. 10

June, 1965

A Factorial Study of the Relationships between Teacher-held Objectives and Student Performance in UICSM High School Mathematics

Hiroshi Ikeda

The material reported here is a dissertation submitted
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Education at the University of Illinois.

The research described in this report
was supported by the
National Science Foundation

A FACTORIAL STUDY OF THE RELATIONSHIPS BETWEEN TEACHER-HELD OBJECTIVES AND STUDENT PERFORMANCE IN UICSM HIGH SCHOOL MATHEMATICS

Hiroshi Ikeda, Ph.D.

Department of Education

University of Illinois, 1965

ABSTRACT

This study was designed to investigate some possible relationships between teacher-held objectives and student performance that appeared in the samples of test items of the new UICSM high school mathematics. Two test booklets were made on the basis of the first five chapters of High School Mathematics, Course 1 by M. Beberman and H. E. Vaughan. They were given to the same Pekin High School students in Illinois before and after the contents of the test had been taught. For each item, most of which consisted of four sub-items, scores were obtained for each student both for pre-test and post-test administrations, and the gains between the two were computed. The sample consisted of 154 9th-grade students. The same test booklets were sent to teachers who were currently using the text, and they were asked to evaluate the suitability of each item for use in achievement tests for their students. Their reactions were assumed to be an indirect indication of their objectives in teaching the UICSM text. The sample consisted of 105 teachers from 70 different schools in 19 states.

Intercorrelations among the 54 items of the two tests were computed for each case of student pre-tests, post-tests, gains, and teacher ratings. Each intercorrelation matrix was factor analyzed by the principal component method and three, two, two, and five factors, respectively, were extracted from these cases. For the case of student pre-tests, the first principal factor indicated a general aptitude for learning mathematics. The second principal factor was, in

part, related to position of the items in the tests, and the third was left unidentified.

For the case of student post-tests, the first principal factor indicated a general achievement in the contents of the given chapters. The second principal factor was considered as a deductive-inductive factor. For student gains, no interesting interpretation of the factors was warranted.

For teacher ratings, the first principal factor was related to the general tendency of each teacher's ratings i.e., the teacher's general response set. Items asking for understanding of the basic mathematical concepts tended to have high coefficients on this factor. The second principal factor was related, in part, to a preference for conventional vs. new mathematics items. The third factor was related to objectives somewhat irrelevant to the text. The fourth factor was related to the objective of algebraic manipulation, and the fifth was left unidentified.

By a canonical type of analysis, the factors from teacher ratings were matched with those from student performance for each of the three cases so that the similarity between the two sets of factor coefficients was maximum. In this analysis, pairs of highly congruent factors were obtained both for the pre-test and for the post-test cases. In the case of student gains, however, no factor was significantly congruent with teacher ratings. However, mean gain scores showed a positive relationship with mean teacher ratings except for a few special items.

FOREWORD

Workers in curriculum evaluation have become increasingly aware of the fact that some of the best efforts at curriculum reform produce superior learning only in the hands of a relatively small group of teachers. Consequently, it would be of little use to learn that curriculum X produced significantly better learning in 50,000 randomly selected students than curriculum Z did in 50,000 comparable students under the tutelage of 2,000 randomly assigned teachers — even if such classically ideal information were available.

School System A should want the best curriculum for its own teachers, who may not be at all representative of the 2,000, or it may plan to hire (or train) teachers who can teach much more to its students than the average learned by the 100,000. Curriculum Z could easily be the best for System A, especially if it succeeds in upgrading its staff. Furthermore, hundreds of school systems around the country may satisfy our description of System A.

Curriculum evaluation for ambitious schools, then, needs to be carried out relative to some characteristics of teachers related to the content of the programs being evaluated. Many studies of general, teacher variables have been carried out, and in a few cases, such variables have been shown relevant to pupil accomplishment in the new curricula.* However, most measures of classroom behavior, personality, and training of teachers seem, a priori, to have little relevance for use in evaluation of new curricula. This does not deny the value of further studies with general teacher variables, but it does suggest that some approaches to content oriented teacher variables should be developed.

---

* Alpert, R., et al, Psychological factors in mathematics education. SMSG Newsletter, April, 1963, No. 15, 17-24.

Spaulding, Robert L., Achievement, Creativity, and Self-Concept Correlates of Teacher-Pupil Transactions in Elementary Schools. Coop. Res. Proj. No. 1352, University of Illinois, 1963.

FOREWORD (Continued)

The UICSM Mathematics Project has been fortunate in having been able to interest Dr. Ikeda in this general problem. The report which follows represents a preliminary exploration of teacher-held objectives for UICSM first year algebra classes and some relations between these objectives and student achievement. This report, which was submitted as Dr. Ikeda's dissertation, is more detailed and more technical than most previous UICSM Research Reports. However, we have decided to issue his thesis, in its entirety, as a number in this series because of the potential value we see in Ikeda's techniques for other researchers interested in teacher-relative curriculum evaluation.

J. A. Easley, Jr.
UICSM Research Coordinator

To the memory

of my mother and father

ACKNOWLEDGMENTS

I wish to express my gratitude to Dr. Ledyard R. Tucker, Chairman of the doctoral committee, for his advice on research design and analysis, to Dr. J. Thomas Hastings, Dr. Robert E. Stake, Dr. Lloyd G. Humphreys, and Dr. Lawrence M. Stolurow for their valuable suggestions, and to Dr. Max Beberman, Director of UICSM, for his help in test construction and for the use of the UICSM facilities.

A special acknowledgment is offered to Dr. John A. Easley, Jr., Research Coordinator of UICSM, for his help given at every stage of the research from research design through data collection, data analysis, careful readings of the draft, up to the final printing. Without his continuous help, this study would not have been realized.

I am also especially indebted to Mr. James Kraatz for his help in computer programming and operation, and to Miss Donna Davis for her help in scoring of the tests. My thanks are due to Mrs. Nancy Zukas, with assistance of Miss Rose Vanerka, for their laborious typing, and to Mr. Roy Lipschutz for the diagram drawing, and to Mr. Harold Bridgewater for the printing. I am also indebted to persons too numerous to mention who assisted significantly at various points.

I am grateful to the students, teachers, and administrators of Arlington Heights High School and Pekin Community High School whose willing cooperation made possible the development of instruments and collection of data, and to the teachers who cooperated by completing the questionnaires, and finally to the University of Illinois Committee on School Mathematics for financial support of my work on the project.

Urbana, Illinois
June, 1965

Hiroshi Ikeda

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iv

TABLE OF CONTENTS v

LIST OF TABLES vii

LIST OF FIGURES ix

CHAPTER I. INTRODUCTION 1

1. Teacher Judgments of Achievement Test Items 2

2. Relationships between Teacher Judgments and Student Performance on Test Items 5

CHAPTER II. THEORY AND HYPOTHESES 6

1. Student Performance and Teacher Preference for Items. 6

2. Post-test Achievement vs. Gained Achievement 7

3. A Mathematical Model Relating Two Factorial Structures 9

A Mathematical Rationale for Matching Two Sets of Factors 15

Summary of the Computational Procedures 29

CHAPTER III. THE COLLECTION OF DATA 31

1. The Construction of Achievement Tests 31

2. The Sample of Students and Test Administration . . . 37

3. The Construction of a Questionnaire for Teachers . . 39

4. The Sample of Teachers 40

TABLE OF CONTENTS (Continued)

Page

CHAPTER IV. RESULTS 45

1. Means and Standard Deviations of Item Data 45

Means and Standard Deviations of the Student Scores for Each Item 45

Means and Standard Deviations of the Teacher Ratings for Each Item 47

2. Intercorrelations among Test Items 62

Intercorrelations of Student Scores among Test Items 62

Intercorrelations of Teacher Ratings among Test Items 62

3. Factor Analysis of Intercorrelations 63

CHAPTER V. DISCUSSION 82

1. Interpretations of the Results 82

Mean Student Performance 82

Mean Teacher Ratings 86

Principal Factors for Student Performance 93

Principal Factors for Teacher Ratings 97

Congruent Factors between Student Performance and Teacher Ratings 99

2. Problems and Suggestions for Future Study 102

CHAPTER VI. SUMMARY AND CONCLUSIONS 105

BIBLIOGRAPHY 109

VITA 112

APPENDIX A. TABLES OF INTERCORRELATIONS AMONG TEST ITEMS

APPENDIX B. THE QUESTIONNAIRE FOR TEACHERS

APPENDIX C. TEST I

APPENDIX D. TEST II

LIST OF TABLES

Page

Table 1. Classification of Items with Respect to Answer Forms . . . 35

Table 2. Classification of Items with Respect to Required Abilities 35

Table 3. Classification of Items with Respect to the Degree of Relevance to the Text 36

Table 4. Geographical Distribution of the Teachers and Schools from Which the Questionnaires Were Collected 42

Table 5. Distribution of the Years of Experience of Teaching Mathematics 43

Table 6. Distribution of the Years of Experience of Teaching Units I and II of the UICSM Old Text 43

Table 7. Distribution of the Number of Courses Taken in Which the Contents of the UICSM Curriculum Were Studied 44

Table 8. Means and Standard Deviations of Pre-test, Post-test, and Gain Scores for Each Item 51

Table 9. Distribution of Mean Scores of Pre-test Items 53

Table 10. Distribution of Mean Scores of Post-test Items 53

Table 11. Distribution of Mean Gain Scores for Items 54

Table 12. Means and Standard Deviations of Teacher Ratings for Each Item 56

Table 13. Distribution of Mean Teacher Ratings of Items 57

Table 14. Principal Factor Coefficients of Student Tests 69

Table 15. Principal Factor Coefficients of Teacher Ratings 71

Table 16. Principal-Factor Similarities between Student Performance and Teacher Ratings 73

LIST OF TABLES (Continued)

Page

Table 17. Cross-Products of the Principal-Factor Similarities . . . 74

Table 18. Transformation Matrices and Latent Roots 75

Table 19. Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Pre-test Performance 76

Table 20. Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Post-test Performance 78

Table 21. Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Gains 79

Table 22. Mean Student Gains and Teacher Ratings for Different Types of Items 85

Table 23. Ranks of Abilities Stressed by the UICSM Teachers 89

Table 24. The Preferred Proportions of Items of Different Kinds to be Included in a Test for Grading 89

LIST OF FIGURES

		Page
Figure 1.	Bivariate Distribution of Mean Scores on Pre- and Post-test Items	55
Figure 2.	Bivariate Distribution of Means on Teacher Ratings and Student Pre-test Scores	58
Figure 3.	Bivariate Distribution of Means on Teacher Ratings and Student Post-test Scores	59
Figure 4.	Bivariate Distribution of Means on Teacher Ratings and Student Gain Scores	60
Figure 5.	Relationship between Means and Standard Deviations of Teacher Ratings for Items	61
Figure 6.	Matching of the Normalized Congruent-Factor Coefficients.	80

CHAPTER I

INTRODUCTION

This study was designed to investigate some possible relationships between teacher-held objectives and student achievement that appeared in responses to sample tests of the new UICSM high school mathematics.

More than thirteen years have passed since the University of Illinois Committee on School Mathematics (UICSM) started to develop new high school mathematics curricula. The unique ideas developed by UICSM on teaching modern mathematics at the high school level have been widely recognized, and a great number of schools are now offering courses based on this curriculum.

To teach mathematics in the way expected by UICSM authors, however, is not easy for teachers who were trained in traditional or conventional ways, and sometimes the ideas involved in the new curriculum arouse confusions and controversies among high school teachers. Although the UICSM provides a training program for high school teachers, the purpose of this program is not always perfectly attained. It is reported by class-room observers that high school teachers trained in UICSM institutes teach mathematics in different ways, and not all of them teach the new curriculum as it should be taught. The high school teachers have their own interpretations of and objectives for the new materials, which are not necessarily conformable to what the UICSM really intends. The importance of studying what kinds of objectives the high school teachers actually have has been often pointed out, since the success of teaching the new mathematics greatly depends on what the teachers think of it and how they react to it. The study was stimulated by this kind of demand.

When we refer to 'teacher-held objectives' for mathematics, it does not mean that the objectives are always explicitly expressed or consciously recognized by every teacher. Some of the objectives might be clear and could be defined by explicit words. Some of the objectives, however, might be unclear and vague and teachers might be unable to express them in a formal statement. Such objectives might only be expressed implicitly, and even teachers themselves might not be conscious of the objectives they hold. This implies that direct questions do not always uncover such hidden objectives. Furthermore, researchers themselves do not know precisely what kinds of questions should be asked of the teachers. In this study teacher-held objectives are examined in indirect ways. Prior to the analysis, we did not assume any definite categories or structure of objectives such as Bloom (1956) and Guilford (1956) have suggested, nor any rationale for the use of objectives such as Mager (1962) has proposed. The objectives we were interested in were those that arise specifically from the interactions of the new UICSM text and the teachers who use it.

1. Teacher Judgments of Achievement Test Items

One approach taken in this study is to examine the teachers' judgments of small groups of test items, many of which are purported to test the abilities students attained in a new mathematics course. Some of the test items are less related to these abilities. Test items were presented for evaluation to high school teachers who were using the latest edition of the UICSM first course text. Most of the items were presented in groups of four which dealt with the same topic and had the same form. A few items were presented individually. The teachers were asked to judge the suitability of each group of test items for use in achievement tests in their own classes. In the rest of this report these groups of test items (usually consisting of four but sometimes single test items) are referred to as the "items" of the instruments used. They are in fact items on the teachers' questionnaire and are the smallest units of analysis for the student achievement tests, since responses to sub-items were not analyzed.

Teachers' judgments of such items are assumed to be consequence of their objectives for the new mathematics course.

An aim of the achievement tests is to measure the abilities that students have acquired during the study of given part of the material. School teachers would teach the subject materials in their own way, as they believe best. After some teaching of the material, the teachers would be likely to know how much their students have progressed by their teaching. In the construction of achievement tests to measure the students' progress, we assume they would be likely to prepare test items that are closely related to what was stressed during their teaching. It is believed that, when sets of ready-made test items are presented to the teachers, they tend to express agreement or disagreement with them according to their own values in teaching the subject. It would be reasonalbe, then, to assume that the teachers' reactions to the items are an indication of their objectives for teaching the subject.

We are not asking here for every detail of the objectives held by individual teachers. To study the case-by-case objectives in every detail is so complex that it is beyond the scope of the present study. In a rough sense, however, objectives held commonly by all of the teachers, or at least, by some group of the teachers could be identified. What we are interested in are such common objectives that appear among all, or some, groups of the teachers when they evaluate samples of achievement test items. We are interested in what kinds of objectives and how many different objectives, at the minimum, would be necessary to account for a set of teachers' reactions.

This notion suggests the application of the factor analysis model developed in the field of psychometrics. The model of factor analysis is useful especially when the aim of research has a somewhat exploratory nature. It has been widely applied to educational and psychological researches. It has been used to investigate basic mental abilities, personality traits, social attitudes, and other human attributes underlying sets of educational and psychological measures. Detailed review of past studies by factor analysis is not the main purpose of this report. Some of these studies, however, should at least be mentioned. These include, for example, Thurstone (1938), Thurstone and Thurstone (1941), French (1951), and Guilford (1956) in the area of mental abilities; Eysenck (1953), Guilford and Zimmerman (1956), and Cattell (1957) in the area of personality studies; Borgatta, Cottrell and Meyer (1956) and Schuessler and Driver (1956) in the area of social psychology.

In a study closely related to our present one, L. R. Tucker (1962) analysed the relevance judgment on the test items of "Developed Abilities of the Social Studies," which were to be given to college applicants. 225 test questions were rated by seventeen experts, and two different kinds of view points were found by the factor analysis of intercorrelations among the raters. The raters were grouped into two different kinds. Group A was a group that emphasized problem-solving ability for both secondary school instruction and for examining students. Group B was a group that emphasized the development of a facility to organize material and to express generalized conclusions effectively.

J. W. French (1962) analysed the reader disagreement that appeared in the scoring of an essay type test. Fifty-three participants consisting of English teachers, social scientists, natural scientists, writers, editors, lawyers, and business executives read and graded 300 college freshman essays. The intercorrelations of the graded scores among participants were factor analysed and six factors were extracted. 101 papers out of 300 received all nine different grades, no paper received less than five different grades, and the average correlation was only .31. It was obvious that some raters were stressing one quality and others stressing another.

Both studies suggest to us that a factor analysis of teacher ratings of items may serve as a useful model to account for the different points of view teachers have on a given set of test items. Part of the present study consists of a factor analysis of teacher ratings of test items, although we are not interested in a typology of teachers as were the above mentioned studies. Our concern is rather with the structure of teacher judgments relative to the items themselves.

2. Relationships Between Teacher Judgments and Student Performance on Test Items

In addition to investigating teacher-held objectives, it is also important to see whether or not students who are studying the UICSM materials develop abilities along the lines which the UICSM authors stress. We don't really know what students learn from the new mathematics curriculum. More specifically, we don't know what kinds of abilities are described in a set of items representative of the UICSM text. Also, we don't know what kinds of items are most associated with respect to student individual differences, developed in a new mathematics course. Such information, even at the descriptive level, would be valuable for the future development of new mathematics curricula. These concerns suggest the usefulness of a factor analytic study of the performance of UICSM students.

Furthermore, it would also be valuable to discover if there is any relationship between the objectives UICSM teachers hold and what UICSM students learn. It would be especially interesting if means were available for collecting data on the objectives of a sample of teachers and also on the achievement of the students of each of them. However, the administration of this type of study was judged to be too complex for existing resources. Besides it seemed wise to invest the available resources in a more exploratory type of study. It is necessary to discover some characteristics of the population and the instruments before one can design with confidence a study of casual relationships between teacher-held objectives and student achievement.

Nevertheless, when we have test items which are supposed to measure student abilities on the contents of text materials, it would be very interesting to compare information on the teachers' judgments of the test items with the actual student performance on the same items. The information for student performance would be helpful for the interpretation of the teachers' objectives found from the analysis of their ratings on items. Also the information from the analysis of teacher ratings on items would be helpful in the interpretation of the students' achievement on the same items.

Many of the items provided in this study are purported to test the abilities that students should attain in the UICSM first mathematics course, but some of the items are not. As we have hypothesized in the earlier paragraphs, if teachers think the items which are stressed in the course are important, then we may expect students to perform best on these items. Moreover, the items which the teachers do not think is important should be ones on which students perform less well — if there does exist a relationship between them. This problem can be investigated by the analysis of correlations between mean values of teacher ratings and mean values of student performance.

On the other hand, a factor analysis of the intercorrelations of student scores on items will tell us the students' common abilities which underlie a given set of items. A factor analysis of the intercorrelations of teacher ratings gives us common factors for teachers' objectives. We are interested in knowing both of these structures. Moreover, if we can develop a technique for investigating the mutual relationship between two sets of factorial structures obtained from the same items, we could ask if there are any consistent patterns between the factors underlying teachers' ratings on test items and the factors underlying students' scores on the same items.

In past studies of factor analyses, I was not able to find an analysis of the mutual relationship between factorial structures of such different performances on a set of items as we have in the case of teacher ratings and student performance. It seemed necessary to extend existing techniques in order to carry out such a study. An appropriate technique for analyzing the relationship between factor structures obtained from entirely different populations is developed in the next chapter.

However, there are some other technical problems which it is necessary to mention briefly. When we relate teacher judgments and student performance on a set of items, what is the best measure to describe student performance? Since we cannot assume that students completely lack the abilities that interest us before they receive instruction, it seems most appropriate to look at the gains made between the administrations of items before instruction and after instruction. However, there are problems in interpreting gain scores, discussed in the next chapter, which make it advisable to examine performance on pre-test and post-test administrations separately as well as analyze the gains between them. Also some compromise seems required between the requirements of reliability of an item score and the necessity of measuring many different objectives within the space of a single class. The compromise we have adopted is described in Chapter III.

In this study, we shall analyze the following:

(1) mean values of student performance on pre-test and post-test administrations of items, as well as gains between them, in order to discover on what kinds of items students perform best;

(2) mean values of teachers' ratings in order to see what kinds of items teachers prefer;

(3) the relationship between the mean values obtained in (1) and (2) in order to see if there exists any relationship between teachers' general preferences and student performance;

(4) a factorial structure of student performance to find the minimum dimensions which account for student abilities;

(5) a factorial structure of teacher ratings to find the minimum dimensions which account for individual differences in teachers' objectives; and finally

(6) the relationship between the two sets of factorial structures in order to determine whether there exist any consistent patterns between individual differences in ratings by teachers and those in student abilities.

CHAPTER II

THEORY AND HYPOTHESES

1. Student Performance and Teacher Preference for Items

In this chapter we shall use the generic term "test" to refer to any stimulus material that generates a response that can be scored on some psychologically interesting dimension. When a set of ability tests, small item groups, or even individual test items (generically, a set of tests) is given to a group of students, a factor analysis of the intercorrelations between the test scores would give us information concerning what kinds of basic abilities underlie the set of tests. Also, we can learn the minimum number of different kinds of abilities that have to be assumed in order to account for the intercorrelations between tests over the students. Suppose that we have Test A having a numerical factor with a coefficient (or factor loading) of 0.8 and a verbal factor with a coefficient of 0.2. A student who has a high ability on the numerical factor and a low ability on the verbal factor would be expected to have a higher score on Test A than a student who has a low ability on the numerical factor and a high ability on the verbal factor, even if the two students have the same total on the two factor scores.

On the other hand, when the same set of ability tests is presented to a group of teachers and the degree of the goodness of each test is rated by them, the factor analysis of the intercorrelations between the tests would tell us what kinds of, and the minimum number of, different view points concerning the goodness of the tests that must be assumed to underlie consistent differences in teacher preference in order to account for the intercorrelations between the tests over teachers. Let us suppose that Test A has a high coefficient (or loading) on Factor I, say 0.8, and a low coefficient on Factor II, say 0.2: Under this circumstance, a teacher who emphasizes Factor I and does not emphasize Factor II may be expected to give a high rating score to Test A. On the other hand, a teacher who emphasizes Factor II and does not emphasize Factor I may be expected to give a relatively low rating score to Test A.

Whether these Factors I and II do or do not have any relationship with the numerical factor and the verbal factor which have been found by the analysis of student scores raises a new and interesting question to ask. If the teachers are capable of rating Test A good (or bad) for the reason that it has a high numerical factor and a low verbal factor, and if they rate other tests in the same way, the factorial structure found in the set of students' scores would have some similarity to the structure found in the teachers' rating scores. If the teachers rate the tests without paying any attention, consciously or unconsciously, to the ability factors which the tests are to measure, the factorial structure of the tests obtained from the teachers' ratings would be different from that obtained from the students' performance, since, in such a case, the teachers must have rated different aspects of the tests. One of the purposes of this study is to find whether or not any correspondence exists between two such structures obtained from achievement tests.

2. Post-test Achievement vs. Gained Achievement

The second problem arises from the question: "Do the students' performance scores on a set of achievement tests serve as the best measure to be related to the teachers' rating scores?" An aim of the achievement test is to measure the students' progress during the study of the course. 'Achievement,' therefore, does not mean a level of performance on the task given at the end of the course, but it does mean an amount of progress during the course. In this sense, a gain score of the performance at the end of the course from the performance at the beginning of the course would be a better measure of achievement than the performance simply shown at the end of the course. This requires us to give a pre-test and a post-test, and obtain a gain score from them to be used as a measure of achievement of the students during the course. If we can assume no special coaching on the given test, we could use the same test twice — as a pre-test one time and a post-test the other. However, some careful direction to the students would be necessary at the pre-test time, because they may be unfamiliar with the given material and may become too frustrated and discouraged to attempt the items.

As for the test, a high average gain score between pre- and post-tests is an indication that the tests are measuring well the students' achievement attained in the course. A low gain score may come from several sources. Some tests might be too difficult for students both on pre- and post-tests. If a test is such that its content is not stressed in the course, the gain scores might be low. Also some tests might be so easy that the students do well at both times, and consequently the gain scores might be low. The students might have already known the content of the tests, or some tests might be correctly answered without any knowledge learned in the course. In other cases, the content of the tests might be irrelevant to what the students studied in the course. The gain scores would be low also, even though the difficulty of the tests is moderate at both times. In any case, a low gain score on a test indicates that it does not serve as a good achievement test. Even though most students do well at the end of the course, this does not necessarily mean either that the tests are good or that the teacher's teaching was effective, unless the average gain score is high. Practically, a test having a low pre-test score and a high post-test score tends to be a good achievement test in this sense.

In the previous section, we discussed the problem of looking for the factorial structure of student performance in the post-test, the factorial structure of teacher judgment, and the relationship between the two structures. The same discussion is possible, replacing the post-test performance of the students by the gain scores. Do teachers judge the test having a high gain score to be good? Is there any relationship between the teachers' judgments and the gain scores? To look for the relationship, if any, between the factorial structure of the gain scores and that of the teachers' rating scores is thus the second and main purpose of this study. However, separate factor analyses were planned for the pre-test data, post-test data, and gain scores data in order to avoid problems resultant from overlapping specific factors in the pre-test and post-test data.

Gulliksen's remarks on 'intrinsic vaildity' are appropriate here, since our attempt is quite parallel to what he suggested.

> ....It would be my hope then that the future development of aptitude and achievement testing will be in the directions of greater emphasis on a search for validities that may be fundamental and lasting as contrasted with those that are likely to be fortuitous or transient. It seems that this concept might be denoted by the term 'intrinsic validity' — intrinsic content validity for achievement tests and intrinsic correlational validity for aptitude tests. .... As far as I can see, we have in the achievement testing field the judgment of the experts to rely upon, but we can do much better and more intensive jobs of checking these expert judgments. If judgments are obtained from a number of persons, the techniques of factor analysis can give some idea of the complexity of the system with which the experts are really dealing. Furthermore, a more intensive use of pre-training and post-training tests would probably be of very great value in sharpening the judgments of the experts regarding the relationships among different types of content (Gulliksen (1950, p. 516)).

3. A Mathematical Model Relating Two Factorial Structures

The investigation of a mathematical relationship between two factorial structures belongs to a branch of canonical analysis. The applicability of canonical analysis to the field of educational research was first proposed by Hotelling in his "The most predictable criterion (1935)." Since then several others have supported the applicability of this approach (e.g., Thomson (1947), Burt (1948), Bartlett (1948), Horst (1961), etc.).

In 1948 M. S. Bartlett proposed a concept of internal and external factor analysis. To take his example, suppose that a set of mental tests and a set of physical measurements are made on the same group of persons. The internal analysis of the mental scores will yield the ordinary factor analysis, and also the internal analysis of the physical measurements will yield a factor analysis of physical structure. The mutual relations (if any) of the two groups of measurements would be examined by means of an external factor analysis. Bartlett's method of external factor analysis is a technique for matching the underlying factors for two sets of variates, the mental scores and the physical measures. The factor matrices for both sets of variates are simultaneously related by orthogonal transformations until a factor from one set is maximally correlated with a factor from the other set, identifying the first factor pair. These factors are held fixed while the second pair is identified, and so forth.

As for the stability of factors over different batteries of tests, L. R. Tucker developed "an inter-battery method of factor analysis (1958)." Given two test batteries, postulated to depend on the same common factors, but not parallel tests, factors that are common to the two batteries are determined from the correlations of the tests in one battery with the test in the other battery. Gibson (1960a, 1960b, 1961) also expanded Tucker's method.

Both Bartlett's and Tucker's methods are closely related to our problem, as far as the mathematical technique is concerned. But they are different from our problem in the respect that the former models treat the case in which the same subjects take different tests or the same tests in different conditions, while the latter has to treat the case in which different subjects respond to the same tests in different ways. This type of problem has been treated in the context of factorial invariance in which how much a factorial structure found in one study is invariant for another study is being investigated. The studies by Tucker (1951), Leyden (1953), Barlow and Burt (1954), Zachert and Friedman (1953), Wrigley and Neuhaus (1955), Meredith (1964a, 1964b), and Pinneau and Newhouse (1964) are examples in this line.

The general idea which we will use in this study is the following. Suppose that n items are given to N_S students to work, and let the standard score of the i_S-th student for the j-th item be z_{ji_S}. The score, z_{ji_S}, is assumed to be determined by a composite of several basic abilities, so-called, factors. If m_S is the number of the basic factors which are commonly involved in more than one item, z_{ji_S} is expressed by an equation

$$z_{ji_S} = a_{j1_S} x_{1_S i_S} + a_{j2_S} x_{2_S i_S} + \dots + a_{jm_S} x_{m_S i_S} + u_{ji_S}$$

$$(i_S = 1_S, 2_S, \dots, N_S \, ; \, j = 1, 2, \dots, n)$$

The coefficients $a_{j1_S}, a_{j2_S}, \dots, a_{jm_S}$ express the contribution of item j to each of the factors $1_S, 2_S, \dots, m_S$, respectively. They are often called the common factor coefficients or the common factor loadings. In this study, we shall use the term 'factor coefficients' instead of 'factor loadings'. If the coefficient a_{j1_S} is large and other coefficients are approximately zero, one common factor 1_S is dominant on the item j. The coefficients are invariant over the students, if the item is specified, since they are item attributes.

The variables $x_{1_S i_S}, x_{2_S i_S}, \dots, x_{m_S i_S}$ refer to the amount of each basic ability that the student i_S has, and they are often called the factor scores of the student i_S. If $x_{1_S i_S}$ is great, the student i_S is considered to have high ability in the factor 1_S. The factor scores are invariant over the given set of items, if the student is specified, since they are student attributes. x's are assumed to be standardized so that the mean item is 0 and the standard deviation is 1.

The term u_{ji_S} is a function of the item and the student. It varies for each item and student and it is called the uniqueness of student i_S for item j.

A similar model holds for the teacher ratings. If we let the rating of teacher i_T on the item j be z_{ji_T}, the z_{ji_T} is assumed to be a composite of several basic objectives held by the teacher and it is expressed by an equation

$$z_{ji_T} = a_{j1_T} x_{1_T i_T} + a_{j2_T} x_{2_T i_T} + \dots + a_{jm_T} x_{m_T i_T} + u_{ji_T}$$

$(i_T = 1_T, 2_T, \ldots, N_T; j = 1, 2, \ldots, n)$

where $a_{j1_T}, a_{j2_T}, \ldots, a_{jm_T}$ are the common factor coefficients for the item j and $x_{1_T i_T}, x_{2_T i_T}, \ldots, x_{m_T i_T}$ are the common factor scores of the teacher i_T. The u_{ji_T} is the uniqueness of the rating of teacher i_T on the item j. z's and x's are also standard scores.*

Factor analysis is a method for obtaining the common factor coefficients for each item and sometimes to estimate the common factor scores for each person under some additional assumptions.

However, the coefficients $a_{j1}, a_{j2}, \ldots, a_{jm}$ are not uniquely determined mathematically. It depends on how the reference axes are taken, and they are taken by somewhat arbitrary criteria. This is, so-called, a rotation problem. For example, if an observed score of the person i on the item j is 1.5 and the number of common factors is two, both solutions below satisfy the equation of the model.

$$z_{ji} = a_{j1} x_{1i} + a_{j2} x_{2i} + u_{ji}$$

$$1.5 = (.7)(1.0) + (.1)(1.0) + .7$$

$$1.5 = (.5)(1.6) + (.5)(0.0) + .7$$

There are an indefinite number of solutions that satisfy this condition.

One of the criteria which is most commonly used is the principle of "simple structure." If only one common factor is involved in an item, it is said that the complexity is one. If more than two factors are involved in an item, it is said that the complexity is two, three, four, and so on, according to the number of common factors involved. It is obvious that minimum complexity would be desirable for the description of the variable. The principle of simple structure is a criterion of minimum complexity of the factors for each item.**

---

[*] In the interpretation of the factor structure of teacher ratings, we must keep in mind the effects of discarding the mean ratings. A consequence of factor analysis of intercorrelations is that all factors obtained are dimensions of individual differences. In this model for teacher ratings, over-all agreements in ratings are discarded.

[**] This is not the rigorous statement of the principle of simple structure. Details may be found in Thurstone (1947, p. 335).

The other principle would be to take the reference axes so that the contribution of the first factor over all the items (the sum $a_{11}^2 + a_{21}^2 + \ldots + a_{n1}^2$) is maximum, and that of the second factors (the sum $a_{12}^2 + a_{22}^2 + \ldots + a_{n2}^2$) is the next maximum, and so on. This is so-called the principal-axes solution.

These methods are, however, the methods for rotating axes within a single set of items over one group of subjects. What we are considering here is the determination of axes so as to take account of two sets of items over different groups of subjects, simultaneously.

Suppose that we have taken m_S common factors from student scores for n items and m_T common factors from teacher ratings for the same n items. The m_S and m_T are not necessarily equal. Instead of taking the sum of squares of the first factor coefficients over n items to be maximum for each group of data, as the principal-axes method does, we could take the first reference axis so that the first factor coefficients of n items for the student data are as similar as possible to the first factor coefficients of the same n items for the teacher ratings on the basis of the criterion of similarity — the coefficient of congruence — which will be explained later. The second reference axis also could be taken so that the second factor coefficients of n items for the student data are as similar as possible to the second factor coefficients of the same n items for the teacher ratings, under the additional condition that the second factor coefficients are orthogonal to the first one. This process could be continued until the pairs of factor coefficients are no longer similar to each other. The system of matched factors obtained in this way will be considered to be congruent factors. The first factor obtained from the student scores is maximumly congruent with the first factor obtained from the teacher ratings. The second factor obtained from the student scores is the next maximumly congruent with the second factor obtained from the teacher ratings, and so forth.

The degree of congruence will be defined as the sum of cross-products of the matched unit factor coefficients, where the unit factor coefficients refer to the normalized factor coefficients over n items so that the sum of the squared coefficients is unity. This degree of congruence will be called the "coefficient of congruence". The coefficient of congruence roughly resembles a coefficient of correlation, but it is not exactly the same, because the coefficient of congruence refers to the sum of cross-products of the unit variables, while the coefficient of correlation refers to the sum of cross-products of the deviations from the means with unit variances. The coefficient of congruence varies between -1 and $+1$. $+1$ means that the two matched unit factor coefficients are identical with each other, -1 means that they are identical in the opposite direction, and 0 means that they are unrelated.

Defining a minimum coefficient of congruence, below which the pairs of factors are not considered to be congruent any more, we could obtain r sets of pairs of congruent factors. The r is usually less than either m_S or m_T. The rest of the student factors (m_S - r of them) are considered the factors that are involved in the student scores only and non-congruent with the factors for teacher ratings. The rest of the teacher factors (m_T - r of them) are considered the factors involved in the teacher ratings only and non-congruent with the factors for the student scores.

What we are going to do in this study is to find the congruent factors (if any) involved in both student performance and teacher ratings for a given set of items. If any congruent factors are found, the degree of congruence will be studied. Three kinds of data for student performance will be used. The first one is the scores of items on the pre-test administrations, when the contents of the items had not been taught. The second one is the scores of items on the post-test administrations after the contents of the items had been taught. The last kind of data is the student gains from pre-test to post-test on each item. Factors found from each set of student data will be matched with the factors found from the teacher ratings and the degrees of congruence for each case will be compared.

The rest of this chapter will be used for more technical development of the computational procedure for getting congruent factors. Its main idea is borrowed from Tucker (1951) and Wrigley and Neuhaus (1955).

A Mathematical Rationale for Matching Two Sets of Factors

Now, let

$$Z_T = \begin{bmatrix} z_{11_T} & z_{12_T} & \cdots & z_{1N_T} \\ . & . & z_{ji_T} & . \\ z_{n1_T} & z_{n2_T} & & z_{nN_T} \end{bmatrix} \begin{pmatrix} j = 1, 2, \ldots, n \\ i_T = 1_T, 2_T, \ldots, N_T \end{pmatrix} \quad (1)$$

and

$$Z_S = \begin{bmatrix} z_{11_S} & z_{12_S} & \cdots & z_{1N_S} \\ . & . & z_{ji_S} & . \\ z_{n1_S} & z_{n2_S} & \cdots & z_{nN_S} \end{bmatrix} \begin{pmatrix} j = 1, 2, \ldots, n \\ i_S = 1_S, 2_S, \ldots, N_S \end{pmatrix} \quad (2)$$

where n is the number of test items used,

N_T is the number of teachers who judged the items, and

N_S is the number of students who took the tests.

Z_T is an $n \times N_T$ matrix whose (j, i_T) element is a standard rating of the j-th item judged by the i_T-th teacher.

Z_S is an $n \times N_S$ matrix whose (j, i_S) element is a standard score of the j-th item obtained by the i_S-th student.

When we are talking of a gain score of post-test from pre-test, z_{ji_S} would be the gain score of the j-th item of the i_S-th student. When we are talking of a performance score of the post-test only, the z_{ji_S} would be the performance score of the j-th post-test item of the i_S-th student, and so on. As there is no difference between the kinds of student scores in the mathematical formulation, we shall not differentiate these scores in the following discussion unless the difference has a special significance.

For matrices Z_T and Z_S, the items bearing the same row number j represent identical items, but the difference is that the one is given to the teachers to judge and the other is given to the students to work.

For each row of the matrices Z_T and Z_S, the scores of teachers and students are assumed to be standardized, for convenience, i.e.,

$$\frac{1}{N_T} \sum_{i_T = 1}^{N_T} z_{ji_T} = 0 \quad , \tag{3}$$

$$\frac{1}{N_S} \sum_{i_S = 1}^{N_S} z_{ji_S} = 0 \quad , \tag{4}$$

$$\frac{1}{N_T} \sum_{i_T = 1}^{N_T} z_{ji_T}^2 = 1 \quad , \tag{5}$$

$$\frac{1}{N_S} \sum_{i_S = 1}^{N_S} z_{ji_S}^2 = 1 \quad . \tag{6}$$

Hence, the products of Z_T and its transpose Z_T' divided by N_T yields a matrix of correlation coefficients between items judged by teachers,

$$R_T = \frac{1}{N_T} Z_T Z_T' \quad , \tag{7}$$

Similarly,

$$R_S = \frac{1}{N_S} Z_S Z_S' \quad , \tag{8}$$

which is a matrix of correlation coefficients between items of tests for students. Both R_T and R_S are $n \times n$ symmetric matrices whose diagonal elements stand for a unit variance of each item.

According to the ordinary model of factor analysis, each of the matrices Z_T and Z_S is assumed to be factorized as

$$Z_T = A_T X_T + U_T \quad , \tag{9}$$

and

$$Z_S = A_S X_S + U_S \quad , \tag{10}$$

or Z_T and Z_S are approximated by

$$Z_T \doteq \hat{Z}_T = A_T X_T \quad , \tag{11}$$

and

$$Z_S \doteq \hat{Z}_S = A_S X_S \quad . \tag{12}$$

A_T is an $n \times m_T$ matrix ($m_T \leq n$) which may be called a factor coefficient matrix or simply a factor matrix for teacher judgments on items. X_T is an $m_T \times N_T$ matrix ($m_T \leq N_T$) which is called a factor score matrix for teacher judgments on items. A_S and X_S are $n \times m_S$ ($m_S \leq n$) and $m_S \times N_S$ ($m_S \leq N_S$) matrices for student scores, and called a factor coefficient matrix and a factor score matrix, respectively. Neither m_T nor m_S are greater than the number of items n which is not greater than either N_T or N_S, but m_T is not necessarily equal to m_S. Since the analysis of Z_T and Z_S can be done in parallel, we shall simply refer to Z, A, X, U, and N, for a while, instead of differentiating Z_T and Z_S, A_T and A_S, X_T and X_S, U_T and U_S, N_T and N_S, etc.

A and X could be determined in various ways (see Thurstone, 1947 and Harman, 1960). One way is such that the distance between Z and AX with a lower rank than that of Z is minimized in a least squares sense, as in the development by Eckart and Young (1936). (See also Horst, 1963). U could be regarded as an error part of the approximation.[*]

---

[*] The choice of using the Eckart and Young development corresponds to determination of principal axes factors from the correlation matrix with unities in the diagonal cells rather than communalities in these diagonal cells. Recent work by Tucker (1965) using simulated correlation matrices indicates that principal axes factoring of these correlation (continued to the next page)

According to Eckart and Young, a matrix $\hat{Z}$ is constructed to the desired degree of approximation in a form

$$\hat{Z} = FLX , \tag{13}$$

where F is an n × k orthonormal matrix such that

$$F'F = I_k , \tag{14}$$

which is an identity matrix of order k.

L is a k × k diagonal matrix whose diagonal elements are positive and arranged in descending order of magnitude, and

X is a k × N orthogonal matrix such that

$$\frac{1}{N}XX' = I_k . \tag{15}$$

If we let

$$A = FL , \tag{16}$$

equation (13) becomes

$$\hat{Z} = AX . \tag{17}$$

In our present model, we could start from the intercorrelation matrix

$$R = \frac{1}{N}ZZ' . \tag{18}$$

The components F, L, and X in equation (13) could be determined from the latent roots and vectors of the intercorrelation matrix, R. Since R is a square, symmetric matrix with unit diagonal elements, R may be directly analysed by the ordinary principal component analysis (see Hotelling, 1933, and Harman, 1960).

Let R be approximated by

$$\hat{R} = \frac{1}{N}\hat{Z}\hat{Z}' . \tag{19}$$

---

matrices with unities in the diagonal yields factor solutions closer to input common factors than does factor analysis of these matrices with communality estimates in the diagonal cells.

From (13),

$$\hat{R} = \frac{1}{N}(FLX)(FLX)'$$

$$= \frac{1}{N}(FLXX'L'F') \quad . \qquad (20)$$

Since $\frac{1}{N}XX' = I$ and $L = L'$, then

$$\hat{R} = FL^2F' \quad . \qquad (21)$$

We should notice that L^2 is identical with latent roots of r, and F is a set of unit latent vectors associated with L^2. This could be verified by post-multiplying both sides of (21) by F:

$$\hat{R}F = FL^2F'F = FL^2 \quad . \qquad (22)$$

The product of the matrix $\hat{R}$ and by each of the column vectors of F is a vector which is proportional to that column vector of F. In each resulting column vector, the constant of proportionality is the corresponding element of the diagonal of L^2. This means that F is a set of latent vectors of $\hat{R}$ associated with k latent roots. Furthermore, since k latent roots of $\hat{R}$ are the first k largest roots of R (Harman, 1960), the proportion of the sum of the diagonal elements of L^2 to those of R indicates how closely the matrix $\hat{R}$ (= FL^2F') with rank k ($k \leq n$) approximates the matrix R.

Since $A = FL$ by (16), equation (21) yields

$$\hat{R} = AA' \quad . \qquad (23)$$

Since $F'F = I$ by (14) and L is a diagonal matrix,

$$A'A = L'F'FL$$

$$= L^2 \quad . \qquad (24)$$

Equations (23) and (24) also indicate important properties of matrix A.

Let us go back to our problem. Here we have two sets of data, the one is from a group of teachers and the other is from a group of students. They are expressed

$$\hat{Z}_T = A_T X_T = F_T L_T X_T \, , \tag{25}$$

and

$$\hat{Z}_S = A_S X_S = F_S L_S X_S \, , \tag{26}$$

where

$$A_T = \begin{bmatrix} a_{11_T} & a_{12_T} & \cdot & a_{1m_T} \\ \cdot & & a_{jp_T} & \cdot \\ a_{n1_T} & a_{n2_T} & \cdot & a_{nm_T} \end{bmatrix} \tag{27}$$

$$= \begin{bmatrix} a_{1_T} & a_{2_T} & \cdot & a_{p_T} & \cdot & a_{m_T} \end{bmatrix} \tag{28}$$

$$(j = 1,\ 2,\ .\ .\ ,\ n\ ;\ p_T = 1_T,\ 2_T,\ .\ .\ ,\ m_T),$$

and

$$A_S = \begin{bmatrix} a_{11_S} & a_{12_S} & \cdot & a_{1m_S} \\ \cdot & \cdot & a_{jq_S} & \cdot \\ a_{n1_S} & a_{n2_S} & \cdot & a_{nm_S} \end{bmatrix} \tag{29}$$

$$= \begin{bmatrix} a_{1_S} & a_{2_S} & \cdot & a_{q_S} & \cdot & a_{m_S} \end{bmatrix} \tag{30}$$

$$(j = 1,\ 2,\ .\ .\ ,\ n\ ;\ q_S = 1_S,\ 2_S,\ .\ .\ ,\ m_S).$$

The elements of matrices (28) and (30) represent column vectors obtained by partitioning (27) and (29) for each column, respectively.

Now the question is how similar are the two factor matrices A_T and A_S to each other. If the first factor from the teachers' data is similar to the first factor from the students', the vector a_{1_T} is approximately equal to the vector a_{1_S}. If the second factor from the teachers' data is similar to the second factor from the students', the vector a_{2_T} is approximately equal to the vector a_{2_S}, and so forth.

As a matter of fact, the coordinates in factor space are chosen in somewhat arbitrary way. Hence, a transformed matrix

$$B_T = A_T T_T \tag{31}$$

also gives us a factor matrix having different coordinates. T_T is a transformation matrix. Similarly, we could have a transformed matrix

$$B_S = A_S T_S \tag{32}$$

for students' data. Our problem is to choose factor matrices B_T and B_S so that the similarity between two factor matrices is maximized by setting adequate coordinates.

First of all, an index to express the degree of similarity between two factor matrices has to be determined. When two factor matrices are given by (27) and (29), the degree of similarity of a factor p_T in matrix (27) and a factor q_S in matrix (29) could be defined by

$$h_{p_T q_S} = \frac{\sum_{j=1}^{n} a_{jp_T} a_{jq_S}}{\sqrt{\sum_{j=1}^{n} a^2_{jp_T} \sum_{j=1}^{n} a^2_{jq_S}}} \begin{pmatrix} p_T = 1_T, 2_T, \ldots, m_T \\ q_S = 1_S, 2_S, \ldots, m_S \end{pmatrix} . \tag{33}$$

Using the vector notations of (28) and (30),

$$h_{p_T q_S} = \frac{(a'_{p_T} a_{q_S})}{\sqrt{(a'_{p_T} a_{p_T})(a'_{q_S} a_{q_S})}} . \tag{34}$$

If we express all the $h_{p_T q_S}$'s in a matrix form,

$$H_{TS} = D_T^{-\frac{1}{2}} A'_T A_S D_S^{-\frac{1}{2}} , \tag{35}$$

where the elements of H_{TS} are $h_{p_T q_S}$.

D_T is a diagonal matrix whose diagonal elements are equal to the diagonals of $A'_T A_T$, and

D_S is also a diagonal matrix whose diagonal elements are equal to the diagonals of $A'_S A_S$.

If the column p_T of factor matrix A_T and the column q_S of factor matrix A_S are identical with each other,

$$h_{p_T q_S} = 1 \ . \tag{36}$$

If the two columns are orthogonal,

$$h_{p_T q_S} = 0 \ . \tag{37}$$

$h_{p_T q_S}$ varies from -1 to $+1$.

Such a measure of the degree of similarity is the same as Burt's "unadjusted correlation (1948)", Tucker's "coefficient of congruence (1951)," and Wrigley and Neuhaus's "degree of factor similarity (1955)".

Now, the coefficient of congruence, which we shall use in this study, between two factor matrices B_T ($= A_T T_T$) and B_S($= A_S T_S$) in equations (31) and (32) can be maximized by choosing adequate transformation matrices as follows. By defining a congruent space between two factor spaces B_T and B_S, the first factor of B_T is matched with the first factor of B_S so that the highest coefficient of congruence is obtained. The second factor of B_T is also matched with the second factor of B_S with the second highest degree of congruence holding the orthogonality to the preceding factors, and so forth, until a significant degree of congruence can not be obtained any more. Generally, the number of congruent factors are less than the number of either factors of A_T or of A_S. Non-congruent factors are considered as factors accounting for only one of the factor matrices A_T and A_S.

Now, let us consider two orthonormal matrices F_T and F_S such that $F'_T F_T = I_{m_T}$, $F'_S F_S = I_{m_S}$, as seen in equation (14). As F_T and F_S are considered to be normalized factor matrices of A_T and A_S, respectively, over each column, the coefficient of congruence between two factor matrices A_T and A_S defined in (35) is simply given by

$$H_{TS} = F'_T F_S \ . \tag{38}$$

Suppose that F_T is transformed, post-multiplying by a unit m_T-vector v_{1_T} $(v'_{1_T} v_{1_T} = 1)$, and F_S is also post-multiplied by a unit m_S-vector v_{1_S} $(v'_{1_S} v_{1_S} = 1)$. The coefficient of congruence between the transformed vectors $F_T v_{1_T}$ and $F_S v_{1_S}$ is given by

$$h_{1_T 1_S} = (F_T v_{1_T})'(F_S v_{1_S}) \;, \tag{39}$$

since

$$(F_T v_{1_T})'(F_T v_{1_T}) = 1 \;, \tag{40}$$

and

$$(F_S v_{1_S})'(F_S v_{1_S}) = 1 \;. \tag{41}$$

The maximization of $h_{1_T 1_S}$ by choosing adequate weight vectors v_{1_T} and v_{1_S} under the conditions (40) and (41) is well-known problem of canonical analysis (e.g., Anderson, 1958).

We may define

$$\begin{aligned} J &= h_{1_T 1_S} - \tfrac{1}{2}\lambda\{(F_T v_{1_T})'(F_T v_{1_T}) - 1\} - \tfrac{1}{2}\mu\,\{(F_S v_{1_S})'(F_S v_{1_S}) - 1\} \\ &= v'_{1_T}(F'_T F_S) v_{1_S} - \tfrac{1}{2}\lambda(v'_{1_T} v_{1_T} - 1) - \tfrac{1}{2}\mu\,(v'_{1_S} v_{1_S} - 1) \end{aligned} \tag{42}$$

where λ and μ are Lagrange multipliers.

Differentiating (42) with respect to v_{1_T} and v_{1_S} and setting the derivatives equal to zero,

$$\frac{\partial J}{\partial v_{1_T}} = (F'_T F_S) v_{1_S} - \lambda v_{1_T} = 0 \;, \tag{43}$$

$$\frac{\partial J}{\partial v_{1_S}} = (F'_T F_S)' v_{1_T} - \mu v_{1_S} = 0 \;. \tag{44}$$

Pre-multiplying (43) by v'_{1_T} and noticing that $v'_{1_T} v_{1_T} = 1$,

$$(F_T v_{1_T})'(F_S v_{1_S}) - \lambda = 0 \;. \tag{45}$$

Pre-multiplying (44) by v'_{1_S} and noticing that $v'_{1_S} v_{1_S} = 1$,

$$(F_S v_{1_S})'(F_T v_{1_T}) - \mu = 0 \ . \tag{46}$$

Since $$(F_S v_{1_S})'(F_T v_{1_T}) = (F_T v_{1_T})'(F_S v_{1_S}) \ , \tag{47}$$

then $$\lambda = \mu \tag{48}$$

Multiplying (43) by λ,

$$(F_T' F_S) \lambda v_{1_S} - \lambda^2 v_{1_T} = 0 \ . \tag{49}$$

By (48), (44) yields

$$(F_T' F_S)' v_{1_T} = \lambda v_{1_S} \tag{50}$$

Substituting (50) into (49),

$$(F_T' F_S)(F_T' F_S)' v_{1_T} - \lambda^2 v_{1_T} = 0 \ . \tag{51}$$

If we let

$$G_{TT} = (F_T' F_S)(F_T' F_S)' \ , \tag{52}$$

(51) yields

$$(G_{TT} - \lambda^2 I) v_{1_T} = 0 \ . \tag{53}$$

Multiplying (44) by μ and replacing μ by λ,

$$(F_T' F_S)' \lambda v_{1_T} - \lambda^2 v_{1_S} = 0 \ . \tag{54}$$

Substituting (43) in (54)

$$(F_T' F_S)'(F_T' F_S) v_{1_S} - \lambda^2 v_{1_S} = 0. \tag{55}$$

If we let

$$G_{SS} = (F_T' F_S)'(F_T F_S) \ , \tag{56}$$

(55) yields

$$(G_{SS} - \lambda^2 I) v_{1_S} = 0 \ . \tag{57}$$

To find non-zero vectors v_{1_T} satisfying equation (53) is an ordinary latent root problem, or more often called an "eigenvalue problem", and λ^2 is a positive latent root of G_{TT}. Similar relations hold in equation (57). λ^2 is also a positive latent root of G_{SS} and v_{1_S} is a non-zero vector associated with λ^2. The condition that the equation (53) and (57) have non-trivial solution is that λ^2 is a root of the characteristic equations

$$| G_{TT} - \lambda^2 I | = 0 \ , \tag{58}$$

and

$$| G_{SS} - \lambda^2 I | = 0 \ . \tag{59}$$

If the rank of G_{TT} is r, so is the rank of G_{SS}, and (58) and (59) have the same r distinct positive latent roots unless they are special case. Let $\lambda_1^2, \lambda_2^2, \ldots, \lambda_r^2$ be distinct positive latent roots of G_{TT}, such that $\lambda_1^2 > \lambda_2^2 \ldots > \lambda_r^2$, and let $v_{1_T}, v_{2_T}, \ldots, v_{r_T}$ be non-zero latent vectors associated with these roots respectively. The distinct positive latent roots of G_{SS} are also given by $\lambda_1^2, \lambda_2^2, \ldots, \lambda_r^2$ and let $v_{1_S}, v_{2_S}, \ldots, v_{r_S}$ be non-zero latent vectors associated with these roots of G_{SS} respectively.

If we take the first largest root λ_1^2 and the associated vectors v_{1_T} for G_{TT} and v_{1_S} for G_{SS}, the coefficient of congruence between $F_T v_{1_T}$ and $F_S v_{1_S}$ defined by (39) is maximized.

$$\begin{aligned} h_{1_T 1_S} &= (F_T v_{1_T})'(F_S v_{1_S}) \\ &= v'_{1_T} F'_T F_S v_{1_S} \end{aligned} \tag{60}$$

Since $F'_T F_S v_{1_S} = \lambda_1 v_{1_T}$ by (43) and $v'_{1_T} v_{1_T} = 1$,

$$h_{1_T 1_S} = \lambda_1 \qquad (> 0) \ . \tag{61}$$

If we take the second largest root λ_2^2 and the associated vectors v_{2_T} for G_{TT} and v_{2_S} for G_{SS}, the coefficient of congruence between $F_T v_{2_T}$ and $F_S v_{2_S}$ is given by similar procedure to (61), i.e.,

$$h_{2_T 2_S} = \lambda_2 \qquad (> 0) \ , \tag{62}$$

and it is maximum in the residual space defined by

$$(F'_T F_S) - (F_T v_{1_T})(F_S v_{1_S})' \ .$$

Furthermore, it is known that

$$v'_{1_T} v_{2_S} = 0 \ , \tag{63}$$

$$v'_{2_T} v_{1_S} = 0 \ , \tag{64}$$

and $$h_{1_T 2_S} = h_{2_T 1_S} = 0 \ . \tag{65}$$

(See Harman, 1960; Horst, 1963). This procedure can be continued until the vectors v_{r_T} and v_{r_S} associated with λ_r are extracted.

For simplicity, if we define an r × r diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{bmatrix} \tag{66}$$

and an $m_T \times r$ matrix made of column vectors $v_{1_T}, v_{2_T}, \ldots, v_{r_T}$,

$$V_T = [v_{1_T} \ v_{2_T} \ \ldots \ v_{r_T}]^* , \tag{67}$$

equation (53) is expressed in a form

$$G_{TT}\Lambda^2 = V_T \Lambda^2 \tag{68}$$

where $$V'_T V_T = I_r \tag{69}$$

by the property of latent vectors. Similarly, we can define an $m_S \times r$ matrix of column vectors $v_{1_S}, v_{2_S}, \ldots, v_{r_S}$,

$$V_S = [v_{1_S} \ v_{2_S} \ \ldots \ v_{r_S}]^* . \tag{70}$$

---

[*] In order to satisfy equations (43) and (44), it will sometimes be necessary to adjust the directions of the column vectors of V_T and V_S.

Equation (57) is expressed in a form

$$G_{SS}\Lambda^2 = V_S\Lambda^2 \qquad (71)$$

where $$V'_S V_S = I_r \ . \qquad (72)$$

These Λ^2, V_T and V_S can be obtained by ordinary principal component method for G_{TT} and G_{SS} (Hotelling, 1933; Harman, 1960).

By the way, as seen in the equations (43), (44), and (48),

$$(F'_T F_S)V_S = V_T\Lambda \ , \qquad (73)$$

and $$(F'_T F_S)'V_T = V_S\Lambda \ . \qquad (74)$$

These equations could be used for solving the one set of vectors when the other set of vectors is given. For instance,

$$V_T = (F'_T F_S)V_S\Lambda^{-1} \qquad (75)$$

when V_S is given, and

$$V_S = (F'_T F_S)'V_T\Lambda^{-1} \qquad (76)$$

when V_T is given. Since the V_T can be obtained by solving the latent vectors of G_{TT} and the V_S can be independently obtained by solving the latent vectors of G_{SS}, the equations (75) and (76) can be used for checking computation.

Thus, by the transformations of F_T by V_T and F_S by V_S, we can define new factor matrices B_T and B_S in which the first factor of B_T is maximal matched with the first factor of B_S, the second factor of B_T is next maximal matched with the second factor of B_S holding the orthogonality to the first factors, and so on, until the last at the r-th factor of B_T is least maximal matched with the r-th factor of B_S holding their orthogonality with all other factors. These new matrices are expressed by

$$B_T = F_T V_T \ , \qquad (77)$$

and $$B_S = F_S V_S \ . \qquad (78)$$

The coefficients of congruence between factors in different sets are given by a matrix

$$H_{TS} = B'_T B_S = \Lambda \ , \qquad (79)$$

whose diagonal elements indicate the coefficients of congruence between matched factors and whose off-diagonal elements indicate the coefficients of congruence between unmatched factors. The coefficients of congruence between unmatched factors are supposed to be zero within a rounding error. As the diagonals of $B'_T B_S$ are arranged in descending order of magnitude, the last parts of the diagonals might be small and negligible. Only the first several factors having large coefficients of congruence, say over 0.9, could be considered as 'congruent' factors and the rest as 'non-congruent' factors.

The cross-product of matrix B_T with itself is the identity matrix, i.e.,

$$B'_T B_T = I_r \ . \qquad (80)$$

Similarly,

$$B'_S B_S = I_r \ . \qquad (81)$$

Since the sums of squared coefficients of B_T over n items are all one, as seen in equation (80), the B_T is considered as a normalized factor coefficient matrix for teacher ratings. Similarly, the B_S is a normalized factor coefficient matrix for student scores. The normalized factor matrices would be convenient for the comparison between two factorial structures with different units of measurement.

By the way, when A_T and A_S are already given, as seen in the equations (31) and (32), B_T and B_S could be obtained directly from A_T and A_S by transformation matrices T_T and T_S respectively. Since

$$A_T = F_T L_T \ , \qquad (82)$$

and

$$A_S = F_S L_S \ , \qquad (83)$$

by the definition (16), then

$$F_T = A_T L_T^{-1} , \qquad (84)$$

and $$F_S = A_S L_S^{-1} . \qquad (85)$$

Substituting (84) in (77) and (85) in (78).

$$B_T = A_T L_T^{-1} V_T , \qquad (86)$$

and $$B_S = A_S L_S^{-1} V_S . \qquad (87)$$

Hence,

$$T_T = L_T^{-1} V_T , \qquad (88)$$

and $$T_S = L_S^{-1} V_S . \qquad (89)$$

In the next few paragraphs, the computational procedures for determining factor matrices B_T and B_S will be summarized.

Summary of the Computational Procedures

1. Obtain the intercorrelation matrix R_T between n items judged by teachers and the intercorrelation matrix R_S between the same n items performed by students.

2. By the principal component method, find a diagonal matrix L_T^2 consisting of the first m_T significant latent roots of R_T and their associated unit vectors F_T where $F_T'F_T = I_{m_T}$. Similarly, find the first m_S significant latent roots matrix L_S^2 and the associated unit vectors F_S of R_S where $F_S'F_S = I_{m_S}$.

3. If desired, compute $A_T = F_T L_T$ and $A_S = F_S L_S$ which yield principal components of R_T and R_S, respectively, in Hotelling's sense.

4. Determine the positive latent roots matrix Λ^2 and the unit vectors V_T of G_{TT} defined $G_{TT} = (F_T'F_S)(F_T'F_S)'$ by the principal component method.

5. Obtain $V_S = (F_T'F_S)'V_T \Lambda^{-1}$

6. Using the principal component method, find the positive latent roots matrix Λ^2 and the unit vectors V_S of $G_{SS} = (F'_T F_S)'(F'_T F_S)$ reflecting the column vectors of V_S if necessary. Check if Λ^2 and V_S are equal to those found in the steps 4 and 5, respectively.

7. Obtain $V_T = (F'_T F_S) V_S \Lambda^{-1}$ and check if it is equal to the values of V_T found in step 4.

8. Compute the $B_T = F_T V_T$ and $B_S = F_S V_S$. Both B_T and B_S give us the most contruent factor matrices with corresponding columns.

9. Compute $H_{TS} = B'_T B_S$ whose elements are the coefficients of congruence between two sets of factors. Check if $(B'_T B_S)^2$ is equal to the latent roots matrix Λ^2 of G_{TT} (or G_{SS}) obtained in step 4 (or step 6).

CHAPTER III

THE COLLECTION OF DATA

1. The Construction of Achievement Tests

Our work started on the construction of test booklets for students which were based on the first five chapters of High School Mathematics, Course 1 by M. Beberman and H. E. Vaughan (1964).

The contents of the first five chapters of the text are as follows:

Chapter 1	Numerals
Chapter 2	Real Numbers
Chapter 3	Properties of the Real Numbers
Chapter 4	The Language of Algebra
Chapter 5	Operations and Inverses

Forty-three percent of the total pages of the text are in the first five chapters, and these chapters are presumably to be studied within the first semester, even in the slowest class.

Determining the format of the tests posed a problem. As our purpose is to find a factorial structure of a set of tests having different contents, intercorrelation coefficients between these tests have to be computed. To obtain reliable correlation coefficients, however, it would be desirable that each test, in the general sense, has an interval scale whose scores distribute in several ordered categories rather than in the two alternatives of right or wrong answers. At the same time, we want to prepare different kinds of tests in great variety so that no important contents are excluded from the given set of tests. It is a kind of dilemma to require both that long reliable tests and that many kinds of tests are to be given at the same time within a class hour.

As it was estimated that nearly a hundred short questions could be answered within an hour class period, two test booklets were constructed, designed to solve this dilemma. The first booklet, which will be called Test I hereafter, has twenty-five items all but two of which have four sub-items. Test I covers the contents of the first three chapters of the text. As indicated earlier, these groups of sub-items were treated as units, i.e., as the items for teacher judgment and as short tests of student achievement. They are hereafter referred to as items — the items of the two test booklets. The second booklet, which will be called Test II hereafter, has twenty-nine items covering the contents of Chapters Four and Five of the text. Each of these items also has four sub-items except one which is a theorem-proof type of problem. The sub-items within each item are similar to each other in content and controlled by the same instruction. It is assumed that the sub-items of each item measure the same ability, and the differences among them depend not upon the kinds of content but upon the degrees of difficulty.*

Special rules were adopted for scoring the exceptional items. Items No. 5 and No. 6 of Test I were simply scored right-or-wrong, because these items needed more space for presentation, and there was some difficulty in making four questions of the same kind as well. So as to conform to the other items in scoring, a credit of four was given for a correct answer and no credit for an incorrect answer to these questions. Item No. 29 of Test II was also a single question having no sub-items, but the task was to prove a theorem by a step-by-step deduction, and the students' work was graded from zero to four.

---

* This idea would naturally lead us to the concept of radex theory developed by L. Guttman (1954) by which sub-items within the same item are to constitute a simplex structure. The items here, however, were not actually examined from this point of view.

All the rest of the items have four sub-items, to each of which a score of one was given if the answer was correct and zero if the answer was incorrect. Consequently the scores on each item were distributed over the range from zero to four and the resulting scale was treated as if it were an interval scale. Someone might say that the range of the scale is still too narrow for computing a product-moment correlation coefficient, but we preferred to collect many kinds of these very short tests (items) to having longer but fewer tests. There is one practical advantage when the maximum score of each item is four. When the gain score of post-test from pre-test is computed, the maximum gain score will be positive four and the minimum negative four. By adding a constant value of four to every gain score, the range can be changed to from zero to eight without changing the variance and correlation coefficients, and it is punchable in one column of an IBM card.

The items given in the two test booklets are shown in Appendices C and D. For convenience, to refer to an item of Test II in later discussion, one of the consecutive integers from 26 through 54 will be used as shown in parenthesis by the original item numbers of Test II. In Test I the item numbers from 1 to 25 will be used, to refer to the items.

Data on more items were originally collected, especially for Test I, for experimental testing, because the pre-test at the start of instruction of the course was the first such experience for us, and we could not figure out in advance how students would react to such testing. From the experience of experimental testing with three classes at Arlington Heights High School, near Chicago, both the too-easy items and the time-consuming items were deleted, and the number of items was reduced to 25 so that most students could try every item within a class hour even on the pre-test administration. This was particularly necessary in order that the intercorrelations between items located near the end of the booklet would be meaningful. Items in Test I in Appendix C and Test II in Appendix D are the items finally used.

For both Test I and II, some items contain sub-items of the multiple-choice type, some contain completion type items requiring a simple numerical computation, and some require that students have a logically written solution like the proof of a theorem. To increase the variety of items, some of the sub-items are the same as questions in the text except for the given numerical values while others are less closely related to the examples in the text, although the given objectives are the same. Some of the sub-items are not given in the text at all or are based on material given in later chapters.

In Tables 1, 2, and 3, the items are classified with respect to (1) the form of answer required of the students, (2) the abilities required to answer the questions, and (3) the degree of relevance of the content to the text. In Table 1, the answer forms are classified into three categories, (A) multiple-choice form, (B) completion form with simple numerical values or algebraic variables for answers and (C) completion form with some written work. In Table 2, the items are classified according to whether their questions are based on (A) the student's understanding of basic mathematical concepts, (B) the student's computational skill, or (C) the student's ability to apply basic mathematical principles. In Table 3, the items are classified into three categories. If the questions are the same as or a slightly modified form of a question seen in the text, they are classified as A. If the content of the questions is based on the text but some ability of transfer from the text is required, the item is classified as B. If the content of the questions is not treated in the given chapters or irrelevant to the text, the item is classified as C. Briefly, the items of Type A contain questions closely related to the text, the items of Type B contain questions moderately related to the texts and those of Type C have questions least related to the text.

The items in the two tests were based, in part, on a file of test items which the UICSM Mathematics Project had developed as achievement measures for its first course. Some modifications were made to conform to changes in the latest revision of the text. Some new items were created to cover more adequately the chapters in the text. In this process, consultations were held with the project director and several other staff members. However, some items were included which did not fit the text well in order to allow more opportunity for teachers to disapprove items. The classifications described above were made by comparing the items with the examples and exercises in the text. Although they were done subjectively by the author, they were made only after a considerable amount of detailed discussion of them with staff members and experienced UICSM teachers.

TABLE 1

Classification of Items with Respect to Answer Forms

	Frequencies	Item Numbers
A: Multiple-choice	26	1, 5, 6, 12, 15, 16, 19, 20, 21, 24, 25, 26, 28, 32, 35, 36, 37, 40, 41, 42, 43, 44, 45, 46, 49, 50
B: Filling with Numerals or Algebraic Variables	17	2, 3, 4, 7, 8, 10, 11, 13, 14, 17, 18, 27, 38, 39, 51, 52, 53
C: Written Work	11	9, 22, 23, 29, 30, 31, 33, 34, 47, 48, 54
	54	

TABLE 2

Classification of Items with Respect to Required Abilities

	Frequencies	Item Numbers
A: Understanding of Basic Concepts	19	1, 2, 11, 15, 19, 20, 21, 22, 24, 25, 39, 40, 41, 42, 43, 44, 45, 46, 48
B: Computational Skill	16	3, 4, 7, 8, 9, 10, 13, 14, 17, 18, 26, 27, 31, 32, 33, 34
C: Ability of Application of Basic Concepts	19	5, 6, 12, 16, 23, 28, 29, 30, 35, 36, 37, 38, 47, 49, 50, 51, 52, 53, 54
	54	

TABLE 3

Classification of Items with Respect to the Degree of Relevance to the Text

	Frequencies	Item Numbers
A: Closely Related	22	1, 2, 3, 4, 7, 8, 11, 22, 23, 26, 28, 31, 33, 36, 39, 40, 41, 42, 46, 47, 48, 54
B: Moderately Related	20	5, 6, 12, 16, 24, 27, 29, 30, 32, 35, 37, 38, 43, 44, 45, 49, 50, 51, 52, 53
C: Least Related	12	9, 10, 13, 14, 15, 17, 18, 19, 20, 21, 25, 34
	54	

2. The Sample of Students and Test Administration

Test I and Test II were given to all the students who were using the new text, High School Mathematics, Course 1, in the Pekin Community High School in Pekin, Illinois. They were 9th grade students.

Test I was administered as a pre-test on September 11, 1964, within three days after the new semester started. Test I was also administered as a post-test on October 15 within a few days after instruction in the first three chapters was finished. Test II as a pre-test was given on October 16, the next day after Test I was given. Test II as a post-test was given on December 21, when instruction in Chapters Four and Five was finished in most classes. The testing time was sixty minutes for each administration using an ordinary class hour.

The directions to the students at the pre-testing time on both Tests I and II were as follows:

> This test contains a number of questions on mathematics designed to determine what you may already know of some topics you are going to study.
>
> Most of the problems in this test are closely related to what you have studied previously. Therefore, you can often guess what the answer is, although the way in which the problem is written may be new to you.
>
> This test will be used to find out what connections you can see between what you have studied earlier and what you are about to study. However, if you can find no way to solve some of the problems, do not worry. You are not expected to be familiar with all of the questions and this test will not affect your grade.
>
> Try as hard as possible to answer each question so that you will have a high score. When you meet a difficult question and don't really know the correct answer, guess at it, and you may be right.
>
> To spend many minutes in answering one question would not be wise since the total time allowed is limited. You will have until the end of the class hour to work on the test.

> On some questions you will have to write in your answer while on other questions you will have to circle the correct one of several choices.
>
> Read carefully the directions for each question.
>
> You may do scratch work right on the pages of the test.
>
> If you have done all the problems and still have time, you may reread the questions and check your answers again or you may do other work that does not disturb other people.
>
> If you have questions while you are working, raise your hand.

The first few paragraphs read at the pre-testing time were unnecessary at the post-testing time and the directions were simply:

> This test contains a number of questions on mathematics you have studied. Try as hard as possible to answer each question so that you will have a high score. When you meet a difficult question and you don't really know the correct answer, you may guess at it. This test will not affect your grade. To spend many minutes in answering one question would not be wise....

The rest of the directions were the same as those at the pre-test administration.

As mentioned earlier, however, the original experimental booklets having approximately twice as many items as Test I were given to the students of Arlington Heights High School. The same booklets were used for the Pekin High School students at the first pre-test administration. A special direction was given then to cross out the items which were not supposed to be worked. For the rest of the testing, no special direction was necessary.

Excluding the students who were absent on any of the four test days, the total number of students tested was 154. The number of classes was six, and four teachers were teaching these classes, since two classes out of six were taught by the same teacher. The average registered class size was approximately thirty.

3. The Construction of a Questionnaire for Teachers

Our second task was to make a questionnaire asking the teachers who were teaching the new text to judge the suitability of the items in the two tests for use in achievement tests in their own classes. After trying several forms in personal interviews, the questionnaire shown in Appendix B was constructed. In this questionnaire the instructions to the teachers concerned with Test I is as shown. The instruction concerned with Test II was the same except that the underlined words were replaced by the words in parentheses.

> Using the scales below, rate all of the items of the enclosed Test I (Test II). Indicate your rating for each item by placing an 'x' in one of the 10 boxes of the scale corresponding to the item. Mark the box on each scale which indicates how good the item would be, in your opinion, for inclusion in a test to be given at the end of the first three chapters (Chapter 4 and 5) of the new UICSM text for course one.
>
> Every item except No. 5 and No. 6 (No. 29) has four sub-items. You should ignore differences between the sub-items of a given item and rate each item as a whole.
>
> Please, do not omit any items. If you can describe the reasons for your rating briefly, do so in the space provided at the right of the rating scale.

The rating scale used has ten points from zero to nine which permits ratings to be punched in one column of an IBM card. Point 0 is specified as a 'worthless' item, Point 2 as an 'inferior' item, Point 4 as 'good', Point 6 as 'superior', and Point 9 is specified as a 'perfect' item. From past experience, we had learned that teachers are likely to avoid an unfavorable rating to such kind of question, and consequently the distribution of scale scores tended to be negatively skewed. In the present scale, then, the reference words have been slightly forced to the lower level in the range of the scale given so that the average score would locate near the middle of the scale. (Actually it was 5.82.)

In addition to the rating scales, some questions asking about past experiences concerning their teaching and about their education were included to provide a more adequate description of the sample. Two questions about the teacher's own attitude toward teaching UICSM mathematics and preferences concerning a test construction were also included so as to be helpful for the interpretation of the result of factor analysis of test items. These questions are only used for informal comparison and were therefore not intended to represent an exhaustive or systematic measure of attitudes. However, both of them had been tried out by face-to-face administrations of an earlier version of the questionnaire to teachers at Arlington Heights High School. For details of these questions, consult the sample of the questionnaire in Appendix B.

4. The Sample of Teachers

The questionnaire together with copies of Tests I and II were mailed to approximately two hundred teachers who were using the new UICSM text, Course 1. These were all the teachers in the United States who could be located who were using this edition of the text. The number of questionnaires returned in time for our analysis was 105. This sample of teachers came from seventy different schools scattered in nineteen states including four teachers of the Pekin Community High School, Pekin, Illinois, whose classes were used for the testing of students' ability. The geographical distribution of the teachers who cooperated in our study is shown in Table 4. For twenty-four schools out of seventy, two or more teachers from each school replied to the questionnaire, and for the rest of the schools one teacher did.

The questionnaire was sent and collected during January through March, 1965. As the new text was published in September 1964, the questionnaire was sent sufficiently long afterwards that most teachers should have studied the first five chapters of the text. In fact, there was no teacher who reported that his class had not finished the first five chapters at the time when the questionnaire was completed.

Forty-three teachers out of 105 were women. The distribution of the years of experience of teaching mathematics is shown in Table 5. The most experienced teacher had taught mathematics for forty years. For two teachers this was their first experience in teaching mathematics.

The first five chapters of the new UICSM text for Course 1 are closely related to the Units I and II of the old version of the UICSM text published with soft covers. Table 6 shows the distribution of the years of experience of teaching Units I and II, which should be a good indication of the preparatory experience on the contents of the first five chapters of the new UICSM text. For thirty-seven teachers, this was the first experience of teaching with a UICSM text. As seen in Table 7, twenty-two teachers had not taken any course in which the contents of the UICSM curriculum were studied. Most teachers, however, had taken one or two such courses and seven have taken more than four such courses. Sixty-one teachers acquired this training in one or more summer institutes on the UICSM curriculum.

TABLE 4

Geographical Distribution of the Teachers and Schools from Which the Questionnaires Were Collected

States	Number of Schools	Number of Teachers
California	4	7
Colorado	8	10
Hawaii	1	1
Illinois	15	27
Indiana	5	9
Kansas	1	1
Maine	1	1
Massachusetts	9	14
Michigan	3	6
Minnesota	3	5
Missouri	6	7
Nebraska	1	1
Nevada	1	1
New Mexico	2	3
Ohio	4	5
Texas	3	3
Utah	1	1
Wisconsin	1	1
Wyoming	1	2
	70	105

TABLE 5

Distribution of the Years of Experience of Teaching Mathematics

Number of Years	0	1-2	3-5	6-10	11-15	16-20	21-30	31-40	Sum
Number of Teachers	2	17	21	32	18	4	7	4	105

TABLE 6

Distribution of the Years of Experience of Teaching Units I and II of the UICSM Old Text

Number of Years	No Experience	1-2	3-5	6-9	Sum
Number of Teachers	37	36	25	7	105

TABLE 7

Distribution of the Number of Courses Taken in Which the Contents of the UICSM Curriculum Were Studied

Number of Courses	0	1	2	3	4	5	Sum
Number of Teachers	22	40	20	16	5	2	105

CHAPTER IV

RESULTS

1. Means and Standard Deviations of Item Data

Means and Standard Deviations of the Student Scores for Each Item

Table 8 shows the mean scores and the standard deviations of 154 students for each item. Column 1 indicates the consecutive item numbers for Tests I and II. Columns 2, 3 and 4 indicate the means on pre-test, post-test and gain scores, respectively. Columns 5, 6 and 7 indicate the standard deviations of pre-test, post-test, and gain scores, respectively. Since the scores of Items 5 and 6 can only have the values of 4, if correct, and 0, if incorrect, the standard deviations of these items are relatively larger than those of other items.

Tables 9, 10 and 11 show the frequency distributions of mean scores on the 54 items for the pre-test, post-test, and the gain scores, respectively. The item numbers are also shown in these tables so as to indicate which items were easy, which items were difficult, and so on.

As seen in Table 9, a fairly large number of students correctly answered the items given before the instruction began. The mode of the distribution is between 1.0 and 2.0, and the average of scores over all the items is 1.79, which means that, on the average, the students correctly answered about 45% of the sub-items before they studied the content. Test I was easier than Test II. In fact, easy items were collected for Test I because we were afraid of discouraging students by giving unfamiliar questions at the pre-test administration. On the basis of the experience with the pre-test of Test I, more difficult items were included when Test II was made. The average of mean pre-test scores over Test II was 1.64, and over Test I it was 1.96.

As was expected, the distribution of the mean scores from the post-test administrations is negatively skewed as seen in Table 10. The average of post-test mean scores over all the items is 2.68, indicating that the students correctly answered about 67% of the sub-items, on the average, after studying the content. The average of mean post-test scores over Test I is 2.81, and over Test II it is 2.57. Students also did better for Test I than for Test II at the post-test administration.

Looking at the gain from pre-test to post-test in Table 11, 47 items out of 54 have mean gains distributed within a range from 0.00 to 1.50. The average of mean gains is 0.85 over Test I, 0.94 over Test II and 0.90 over all the test items. Students gained more on the harder Test II than they did on Test I. Generally speaking, Test II was a better instrument than Test I, considering the means of the pre-test, post-test, and gain scores.

The relationships between the three kinds of mean scores are shown in Figure 1. The diagonal lines indicate the amount of gains. No item is located on the lower right-hand side of the main diagonal, which means that there is no item with a gain in the negative direction as far as the mean score is concerned. Items 7 and 8 show excessively large gains. Item 34 is too advanced to expect much gain, since exercises like it do not occur until later in the text. Items 11 and 41 are too easy to have much gain. Items 12, 17, 27, 40, and 47 have large gains and they are most desirable items, if the gain is taken as a criterion of a good achievement-test item. Assuming the null hypothesis that no gain was made in the population from which the subjects were drawn, the distribution of sample mean gain scores of an item may be approximated by the t-distribution when the number of subjects is large, even though the distribution of the gain scores of an item in the population is not normal (Hays, 1963, p. 308; Walker and Lev, 1953, p. 143). The hypothesis of no gain in the population cannot be rejected at the 5% level of significance for Items 6, 9, 10, 13, and 19, and at the 1% level for Items 20 and 41 by the two-tailed test (Walker and Lev, 1953, pp. 151ff). All the rest of the items have gains that are significantly different from zero at the 1% level. Examination of the content of items having large or small gains is interesting and important but it will be postponed to the next chapter.

As for the standard deviation of gain scores for each item, it should be noticed that there are general tendencies such that, (1) the standard deviations of gain scores are likely to be greater than either those of pre-test or of post-test scores, (2) items having extremely high or low means have smaller standard deviations than items having means near the middle of the possible range of score distribution.

Means and Standard Deviations of the Teacher Ratings for Each Item

Table 12 shows the means and standard deviations of 105 teachers' ratings for the same items as given to the students. Item numbers in the table correspond to the numbers used for student data in Table 8. The possible range of the rating scale is 0 to 9 and the middle point of the scale is 4.5. (See the sample questionnaire containing the rating scales in Appendix B.) As the mean ratings over all the items is 5.82, it indicates that teachers are likely to rate items toward the favorable direction. The means of the ratings scattered in a range from 4.5 to 7.0 as shown in Table 13. Considering that the possible range is wider for teacher ratings than for student scores, the relative variability of the mean values of teacher ratings is smaller than that of the student scores.

In order to see the mutual relationship between mean values of teacher ratings and student scores, points having two mean values as coordinates are plotted for each item in Figures 2, 3, and 4. These figures show respectively the relationship of means of student pre-test, post-test, and gain scores with teacher ratings. In any case, there seems to be no strong relationship between student scores and teacher ratings. If we compute the correlation coefficients for these bivariate distributions of mean values, −0.13 is obtained for the pre-test, 0.03 for the post-test, and 0.22 for the gain scores. There is a very slight tendency for correlations to increase from pre-test through post-test to gain, but it is non-significant.

Examining details of Figure 4, however, the low correlation coefficient must have been due to a small number of special items such as Items 7, 8, 17, and 18. Items 17 and 18 are rated lowest by teachers and Items 7 and 8 have excessively large gains compared with other items. If we take off these four items as special cases, the correlation coefficient of mean gain scores with mean teacher rating increases from 0.22 to 0.53. There seem to be no special items like these in other pre- and post-test cases. Even if we take off such items as 14, 17, and 18 in Figure 2, or such items as 13, 14, 18, and 34 in Figure 3, the correlation coefficient would not increase as in the case of mean gain scores. It may or may not be recognized by the teachers themselves, but it seems that teachers do tend to evaluate test items with respect to how well they measure student improvement through the course rather than the simple ability to answer questions at the pre- and post-training stages. However, this is not a strong tendency.

It should be borne in mind, in interpreting results, that there was no direct connection between the samples of teachers and students in these analyses. The sample of teachers represent a wide area of the United States while the sample of students are from one local high school in Illinois. The students in the sample were taught by only four teachers, who constitute only a small part of the total sample of teachers. However, it should be worthwhile looking at the relationship between the mean values of these four teachers' ratings in relationship to their students' performance. The numbers in parentheses in Table 12 indicate the mean values of these teachers' ratings. If we compute the correlation coefficient, over 54 items, between the mean ratings of these teachers from Pekin High School and the mean ratings of the entire sample, 0.43 is obtained. Thus, the mean ratings of Pekin High School teachers are positively, but not highly, related to the mean ratings of the whole group of teachers. Items 13, 14, 17 and 18 are rated low by both Pekin teachers and all the teachers. Items 24 and 45 are rated high for both Pekin and all the teachers. Items 15, 16, 23 and 25 are rated high by the Pekin teachers, but they are near the average for the total sample. Items 35 and 37 are rated high by the total sample, but they are rated slightly lower than average by the Pekin teachers. Items 21 and 22 are rated low by the total sample, but they are slightly higher than average by the Pekin teachers.

It is important to see how these items are related to student performance. In general, the correlation coefficients between the mean ratings for the Pekin High School teachers and their students' performance are very low, and no particular relationship was found. The correlation coefficients, over 54 items, of the mean student pre-test, post-test and gain scores with mean ratings of Pekin teachers are, respectively, 0.04, 0.10 and 0.07.

As we have just indicated, in the analysis of the relationship between means of student gain scores and means of total teacher ratings, a positive relationship was found when a few special items were taken out. The peculiarity of these items still holds for the relationship between the mean student gains and the mean ratings of Pekin High School teachers. If these items 7, 8, 17, and 18, are taken out, the correlation coefficient increases from 0.07 to 0.23 but not as much as in the case of the total sample of teachers. However, Items 15 and

16, which are rated high by Pekin teachers, but not by the total sample, had high means for student post-test scores. These items might have been effected by special emphases of Pekin teachers. However, the gains for these items were not exceptionally large.

Thus, low correlations, in general, which have been found in the analyses between the total teacher ratings and the student performance seem not to be due to the fact that we took our teacher sample from a wide area while taking the student sample from one particular school. Therefore, in the following discussion, we shall consider only the relationships between the ratings by the total group of teachers and the student performance. It is assumed that the Pekin teachers and their students are good representatives of the total population under consideration.

It seems worthwhile to look at the standard deviations of teacher ratings on items. A high standard deviation indicates that the dispersion of teachers' evaluations is high and that the items must have some controversial point. A low standard deviation, on the contrary, indicates a high degree of agreement between teacher evaluations, Figure 5 shows the relationship between means and standard deviations of teacher ratings for each item. There is a high negative correlation between means and standard deviations (-0.80). Items 24, 35, 37 and 45 have higher means and lower standard deviations than the other items. It indicates that most teachers rate them as good items with a high degree of agreement. Items 17, 18 and 22, on the other hand, have lower means and higher standard deviations than the other items. Some teachers must have rated them as bad items and some must have rated them as good ones. More discussion considering the contents of these items will be postponed to the next chapter.

TABLE 8

Means and Standard Deviations of Pre-test, Post-test, and Gain Scores for Each Item

(N = 154)

Item Numbers	Means Pre-test	Means Post-test	Means Gain	Standard Deviations Pre-test	Standard Deviations Post-test	Standard Deviations Gain
Test I						
1	2.38	2.86	.49	.91	1.04	1.10
2	3.23	3.83	.60	1.36	.64	1.40
3	2.98	3.79	.81	1.18	.44	1.25
4	2.47	3.44	.96	1.11	.65	1.23
5	1.27	2.42	1.14	1.87	1.96	2.45
6	1.51	1.71	.21*	1.94	1.99	2.50
7	.92	3.56	2.64	.82	.68	1.08
8	.32	3.58	3.26	.64	.73	.99
9	3.21	3.23	.01*	1.04	.95	1.26
10	2.11	2.16	.05*	1.17	1.03	1.53
11	3.54	3.99	.45	.70	.11	.70
12	1.80	3.39	1.59	1.15	1.01	1.45
13	1.75	1.88	.14*	.78	.96	1.09
14	1.03	1.56	.54	.70	.95	1.12
15	1.99	3.27	1.28	1.04	.68	1.23
16	2.59	3.60	1.01	1.10	.67	1.26
17	1.04	2.92	1.88	.97	1.26	1.50
18	.86	1.97	1.11	.97	1.24	1.48
19	3.12	3.13	.01*	.95	.81	1.12
20	2.06	2.30	.24**	.92	.84	1.16
21	1.60	2.08	.47	.97	.86	1.27
22	2.56	3.16	.60	1.45	.95	1.64
23	1.18	1.81	.63	.90	.73	1.07
24	1.94	2.75	.82	.97	.85	1.19
25	1.55	1.96	.41	1.09	.97	1.45
Mean over Test I	1.96	2.81	.85			

* Non-significantly different from zero at the 5% level.

** Non-significantly different from zero at the 1% level.

TABLE 8 (Continued)

Means and Standard Deviations of Pre-test, Post-test, and Gain Scores for Each Item

(N = 154)

Item Numbers	Means			Standard Deviations		
	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
Test II						
26	3.04	3.64	.60	.89	.65	1.09
27	.66	2.75	2.08	1.26	1.45	1.71
28	2.62	3.01	.39	.90	.77	1.09
29	.36	1.56	1.20	.65	.93	1.13
30	1.47	2.01	.55	1.13	1.22	1.42
31	.07	1.43	1.36	.27	1.32	1.34
32	1.10	1.75	.65	.67	.93	1.16
33	1.20	2.23	1.03	.97	1.23	1.36
34	.04	.34	.30	.19	.73	.72
35	2.21	3.27	1.06	1.27	.98	1.30
36	3.37	3.79	.42	.86	.54	.90
37	1.50	2.76	1.26	1.06	1.13	1.32
38	2.08	2.96	.88	1.39	1.40	1.56
39	2.67	3.68	1.01	1.18	.61	1.23
40	.98	2.91	1.93	1.10	1.12	1.37
41	3.71	3.85	.14**	.51	.47	.70
42	2.18	2.74	.57	1.50	1.23	1.47
43	1.96	3.02	1.06	.93	1.06	1.19
44	1.85	2.39	.54	1.06	1.04	1.38
45	1.79	3.01	1.22	.98	.97	1.27
46	2.50	3.01	.51	1.05	.84	1.18
47	1.13	2.77	1.64	1.26	1.12	1.44
48	2.11	3.41	1.30	1.63	1.15	1.69
49	1.00	1.62	.62	.82	.85	1.09
50	1.27	2.31	1.04	.72	1.37	1.38
51	2.16	3.14	.98	1.43	1.18	1.51
52	.53	1.44	.92	.72	1.11	1.23
53	.92	1.78	.86	1.19	1.42	1.73
54	1.01	2.03	1.02	1.40	1.73	1.66
Mean over Test II	1.64	2.57	.94			
Grand Mean	1.79	2.68	.90			

** Non-significantly different from zero at the 1% level.

TABLE 9

Distribution of Mean Scores of Pre-test Items

Means	Frequencies	Item Numbers
0~	4	8, 29, 31, 34
.5~	6	7, 18, 27, 40, 52, 53
1.0~	11	5, 14, 17, 23, 30, 32, 33, 47, 49, 50, 54
1.5~	11	6, 12, 13, 15, 21, 24, 25, 37, 43, 44, 45
2.0~	9	1, 4, 10, 20, 35, 38, 42, 48, 51
2.5~	6	3, 16, 22, 28, 39, 46
3.0~	5	2, 9, 19, 26, 36
3.5~	2	11, 41
	54	

TABLE 10

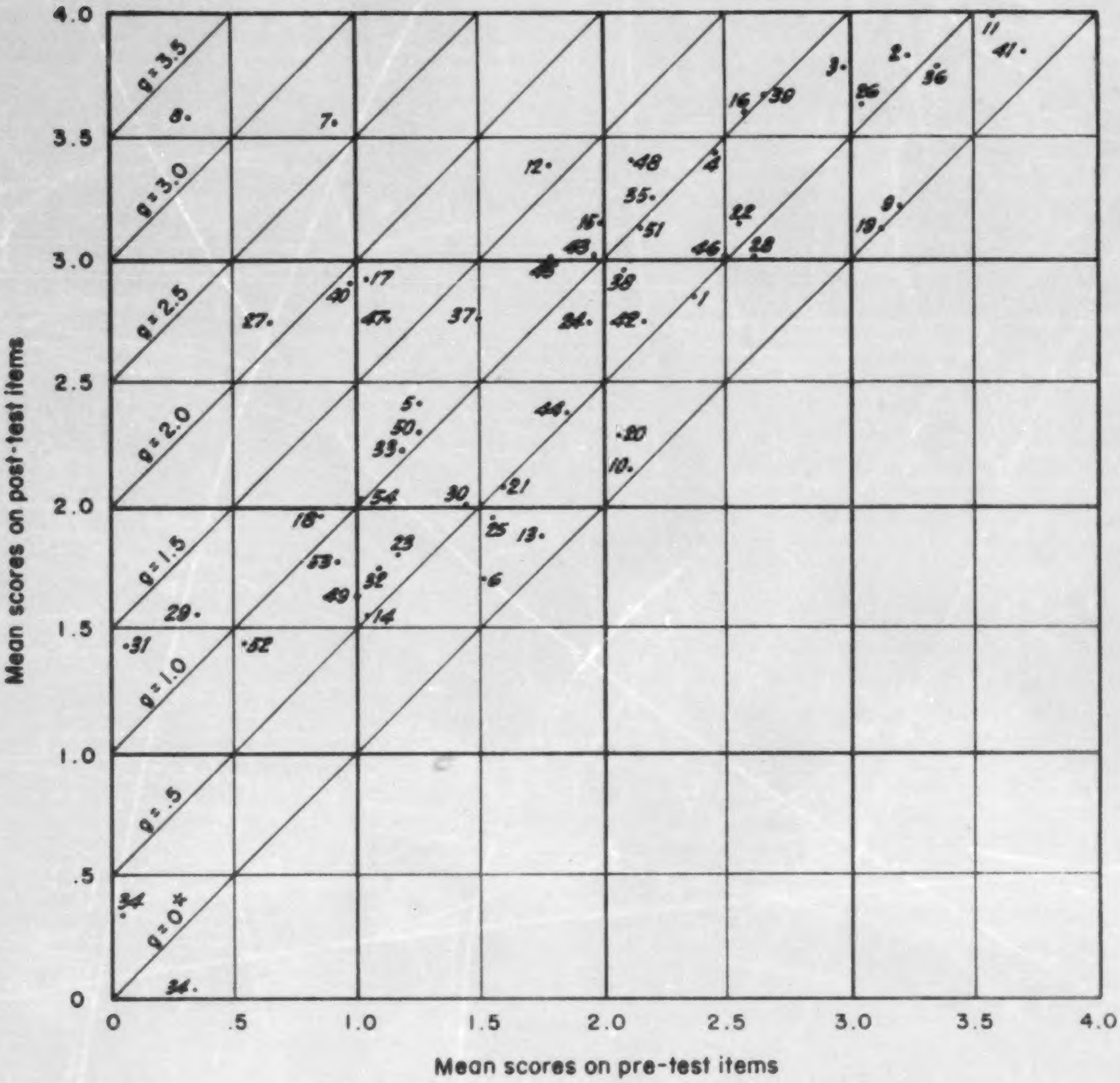
Distribution of Mean Scores of Post-test Items

Means	Frequencies	Item Numbers
0~	1	34
.5~	0	
1.0~	2	31, 52
1.5~	10	6, 13, 14, 18, 23, 25, 29, 32, 49, 53
2.0~	9	5, 10, 20, 21, 30, 33, 44, 50, 54
2.5~	9	1, 17, 24, 27, 37, 38, 40, 42, 47
3.0~	13	4, 9, 12, 15, 19, 22, 28, 35, 43, 45, 46, 48, 51
3.5~	10	2, 3, 7, 8, 11, 16, 26, 36, 39, 41
	54	

TABLE 11

Distribution of Mean Gain Scores for Items

Means	Frequencies	Item Numbers
0~	14	1, 6, 9, 10, 11, 13, 19 20, 21, 25, 28, 34, 36, 41
.5~	18	2, 3, 4, 14, 22, 23, 24, 26, 30 32, 38, 42, 44, 46, 49, 51, 52, 53
1.0~	15	5, 15, 16, 18, 29, 31, 33, 35 37, 39, 43, 45, 48, 50, 54
1.5~	4	12, 17 40, 47
2.0~	1	27
2.5~	1	7
3.0~	1	8
3.5~	0	
	54	



* g = Lines of mean gain scores (= means of post-test scores minus pre-test scores)

FIGURE I

Bivariate distribution of mean scores on pre- and post-test items

TABLE 12

Means and Standard Deviations of Teacher Ratings for Each Item

(N 105)

Item Numbers	Means		Standard Deviations	Item Numbers	Means		Standard Deviations
Test I	Total Sample	Pekin Only*	Total Sample	Test II	Total Sample	Pekin Only	Total Sample
1	5.28	(6.25)	1.98	26	5.62	(5.25)	1.97
2	5.47	(5.75)	2.16	27	6.07	(5.25)	1.74
3	5.30	(6.25)	1.88	28	5.88	(6.00)	1.91
4	5.73	(6.50)	1.79	29	6.11	(5.25)	1.78
5	6.13	(6.00)	1.90	30	6.03	(6.25)	1.75
6	5.57	(6.50)	2.32				
7	5.82	(5.75)	1.83	31	6.27	(6.00)	1.77
8	5.63	(5.75)	1.99	32	6.43	(5.50)	1.66
9	5.33	(4.25)	2.36	33	5.95	(6.00)	1.96
10	5.60	(4.25)	2.36	34	5.79	(5.25)	2.18
				35	6.52	(5.25)	1.60
11	5.84	(5.25)	2.01	36	6.42	(4.75)	1.66
12	6.22	(5.75)	2.28	37	6.74	(5.25)	1.65
13	4.81	(3.25)	2.26	38	6.04	(5.25)	2.03
14	4.86	(3.25)	2.24	39	6.10	(5.00)	2.14
15	5.64	(7.25)	1.99	40	6.35	(6.00)	1.81
16	5.93	(7.00)	2.07				
17	4.73	(3.50)	2.68	41	5.49	(5.25)	2.13
18	4.67	(3.50)	2.64	42	5.80	(5.50)	1.87
19	5.35	(6.25)	2.60	43	6.33	(5.50)	1.74
20	5.56	(5.00)	2.50	44	6.32	(5.75)	1.69
				45	6.53	(6.75)	1.62
21	5.02	(6.00)	2.78	46	5.78	(6.25)	1.94
22	4.81	(6.00)	2.70	47	6.09	(5.00)	2.01
23	5.91	(7.00)	2.18	48	5.87	(6.25)	1.95
24	6.76	(6.75)	1.68	49	6.07	(6.25)	1.97
25	5.86	(7.00)	2.32	50	6.12	(6.50)	1.86
				51	6.08	(5.75)	1.88
				52	6.17	(6.25)	1.87
				53	5.49	(4.50)	2.52
				54	5.90	(6.25)	2.25

	Total Sample	Pekin Only
Grand Mean over Test I	5.51	(5.60)
Grand Mean over Test II =	6.08	(5.66)
Grand Mean over both tests	5.82	(5.63)

* The correlation coefficient of the means for the Pekin teachers with the means for all the teachers is .43.

TABLE 13

Distribution of Mean Teacher Ratings of Items

Means	Frequencies	Item Numbers
4.5~	5	13, 14, 17, 18, 22
5.0~	8	1, 2, 3, 9, 19, 21, 41, 53
5.5~	19	4, 6, 7, 8, 10, 11, 15, 16, 20, 23, 25, 26, 28, 33, 34, 42, 46, 48, 54
6.0~	18	5, 12, 27, 29, 30, 31, 32, 36, 38, 39, 40, 43, 44, 47, 49, 50, 51, 52
6.5~	4	24, 35 37, 45
7.0~	0	
	54	

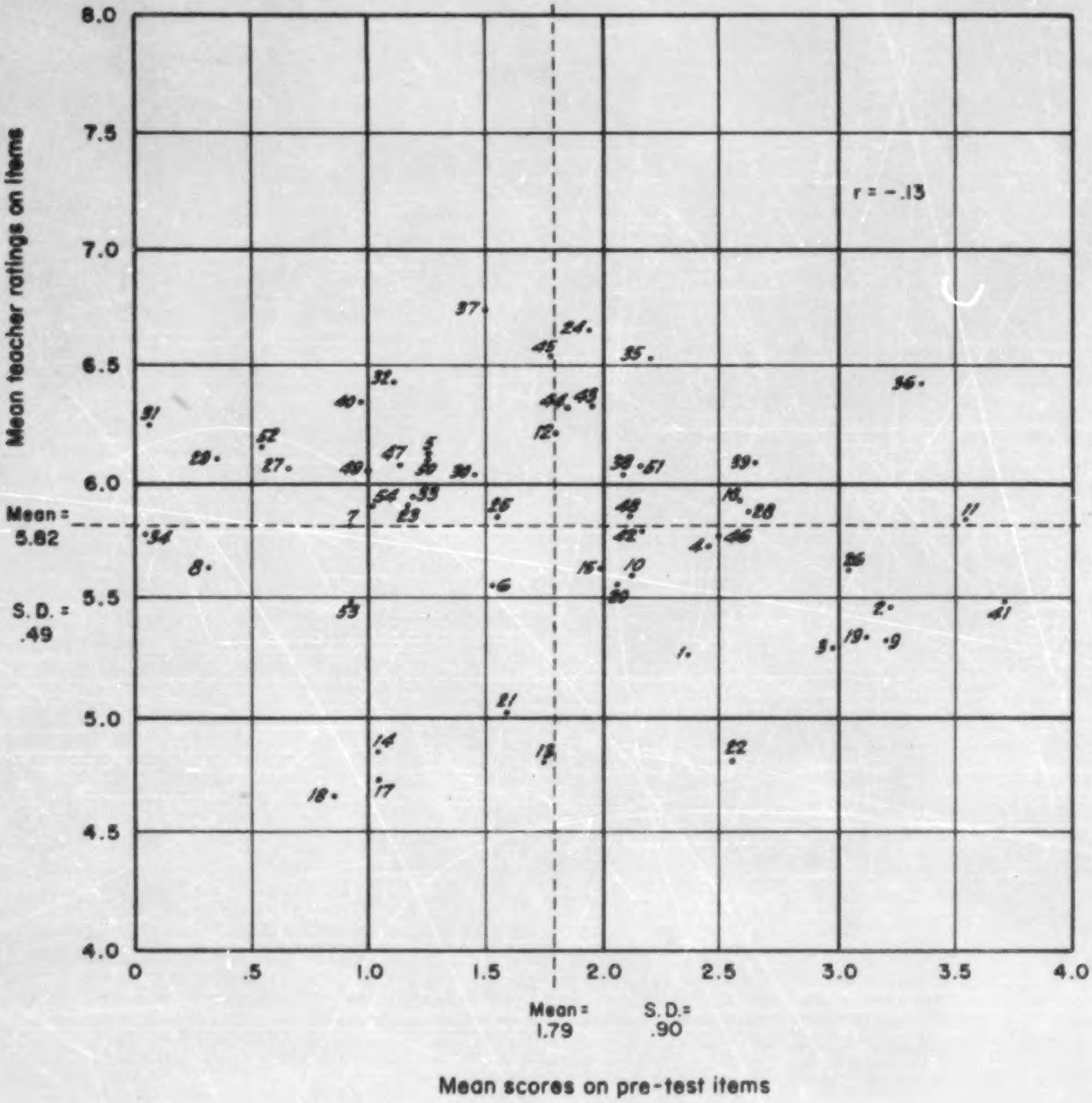


FIGURE 2

Bivariate distribution of means on teacher ratings and student pre-test scores

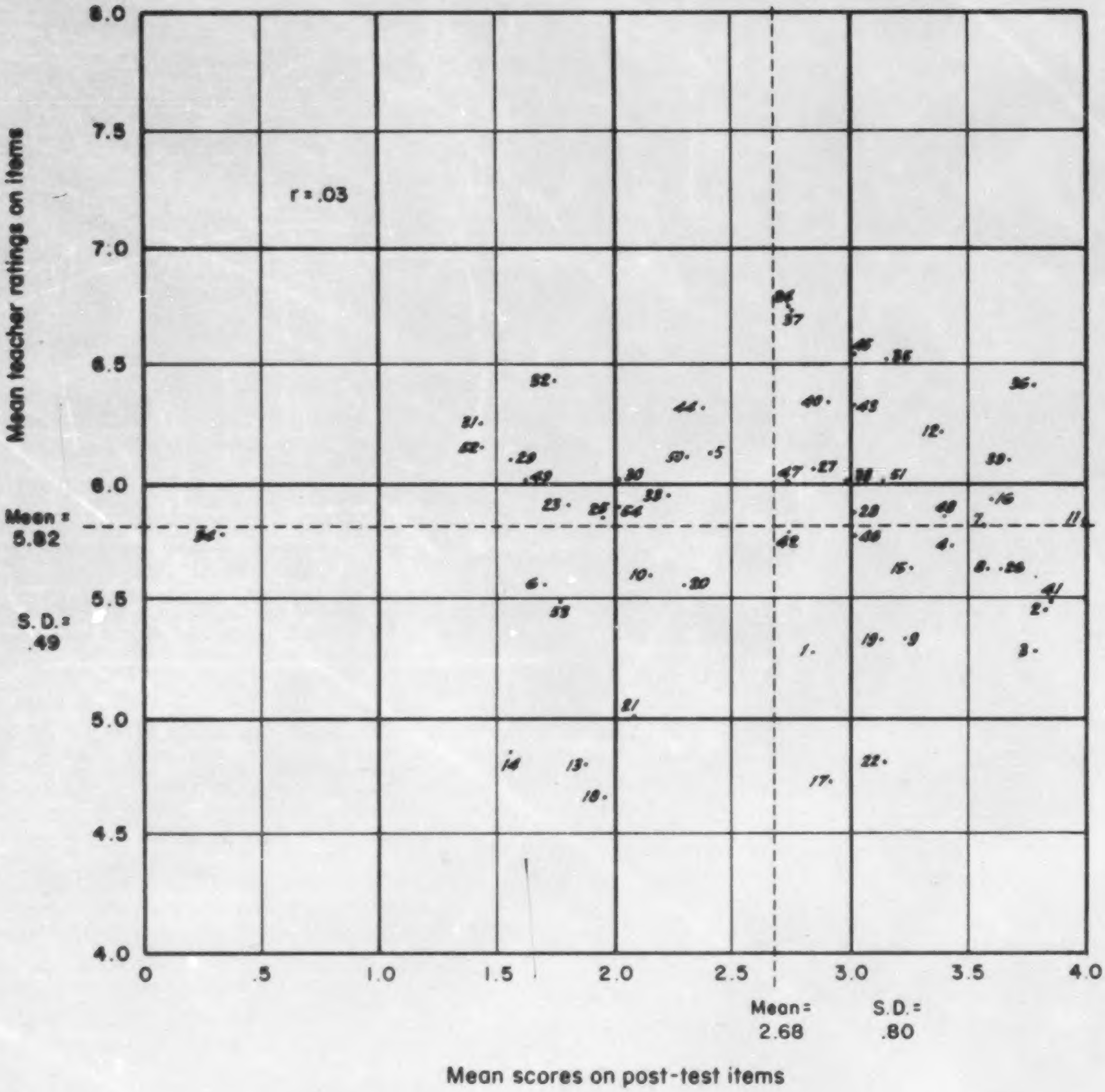


FIGURE 3

Bivariate distribution of means on teacher ratings and student post-test scores

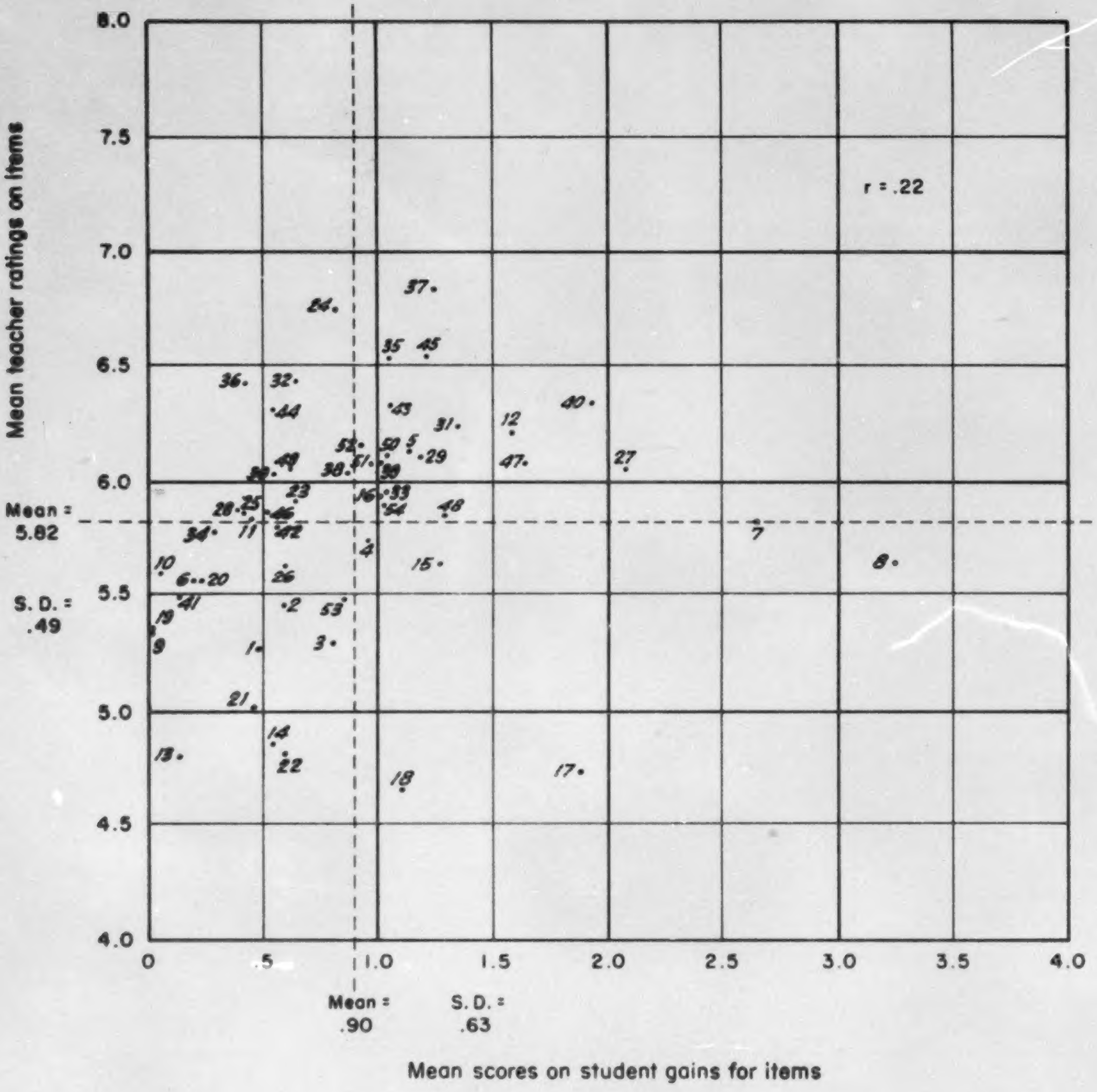


FIGURE 4

Bivariate distribution of means on teacher ratings and student gain scores

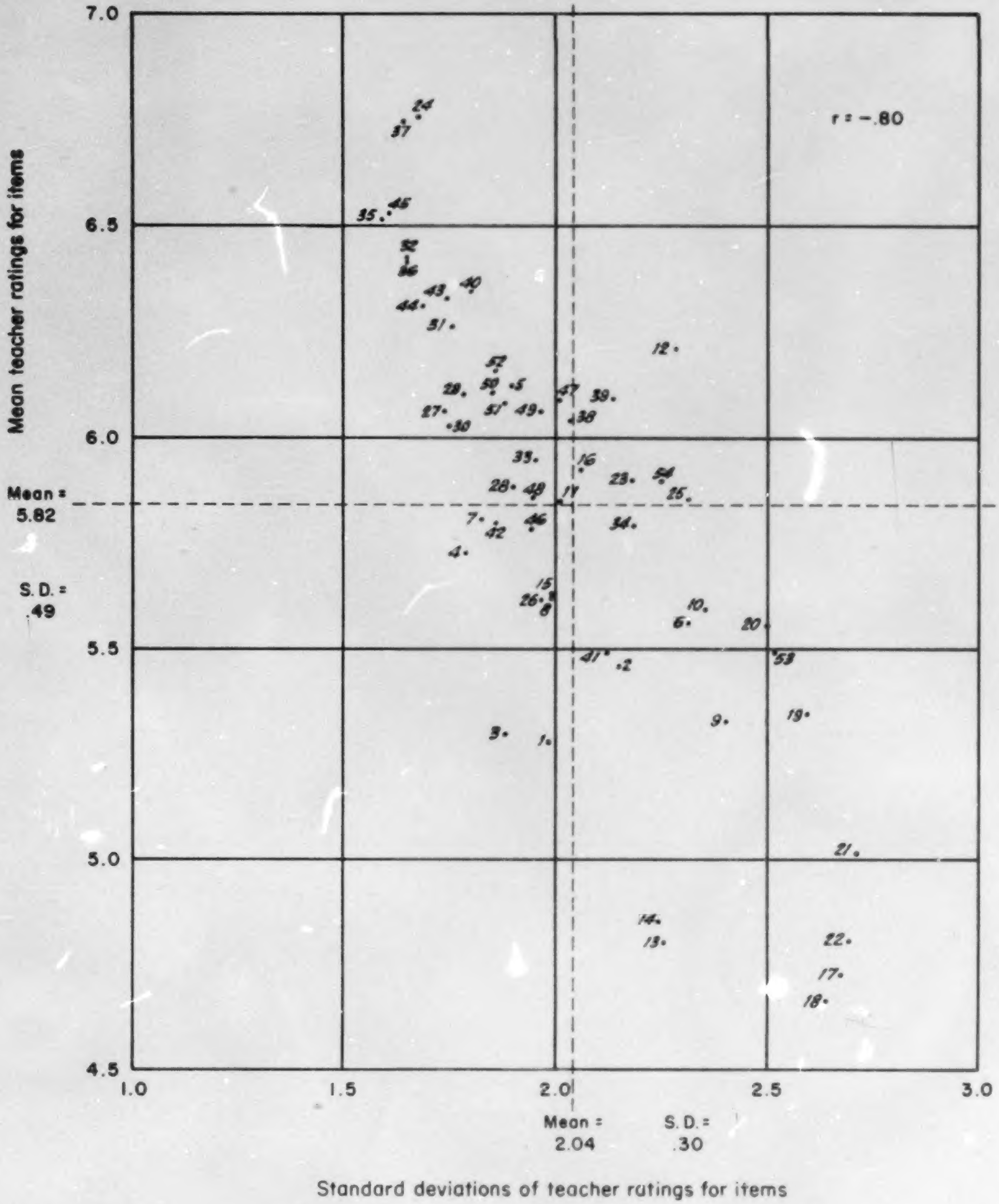


FIGURE 5

Relationship between means and standard deviations of teacher ratings for items

2. Intercorrelations among Test Items

Intercorrelations of Student Scores among Test Items

Tables I, II and III in Appendix A show the correlation coefficients of students' pre-test, post-test, and gain scores respectively, among the items. Generally speaking, the degree of correlations among items is relatively low in all three cases. The correlations which are significantly different from zero at the 5% level are underlined in the tables so as to make it more convenient to find out which items have relatively high correlations with each other.

Comparing the correlations in the cases of pre-test, post-test, and gain scores, the correlations in the pre-test case are generally lower than those in the post-test case, as was expected. The correlations of gain scores, however, tend to be the lowest of the three cases.

In comparing Test I with Test II, it is obvious that the correlations among Test II items are likely to be higher than those among Test I items in the post-test case, but this tendency is not evident in the cases of pre-test and gain scores.

Intercorrelations of Teacher Ratings among Test Items

Table IV in Appendix A shows the correlation coefficients of teacher ratings on the test items. In general, the degree of correlation among items is larger than in any of the three cases of student correlations. In this table, the correlations which are non-significantly different from zero at the 5% level are underlined instead of the significant ones. The correlation coefficients among the items within Test II tend to be higher than the coefficients within Test I. The fact that correlation coefficients among items within Test I are lower than those within Test II indicates that the similarity of Test II items is larger than that of Test I items with respect to the teachers' ratings. A similar relation has been seen for the post-test scores of student data. These results suggest that the items of Test II are seen by teachers as more closely related to the curriculum in the text than Test I items. In fact, as seen in Table 3, Test II items have been constructed so as to be closer to the text than Test I items.

3. Factor Analysis of Intercorrelations

Each of the four kinds of intercorrelation matrices shown in Tables I, II, III, and IV of Appendix A was factor analyzed by the principal component method with unit variances inserted in the diagonals. Three factors were extracted from the correlation matrix among pre-test items. Two factors were extracted from each of the correlation matrices among post-test items, and among the gains. Finally, five factors were extracted from the correlation matrix among teacher ratings. These numbers of factors correspond to sharp breaks in the plots of root size versus root number for the student pre-test, post-test and teachers' rating data. The decision for the gains score was more equivical.

Let F_1, F_2, F_3, and F_4 be normalized latent vectors of R_1, R_2, R_3, and R_4, respectively, where R_1, R_2, R_3 and R_4 denote the intercorrelation matrices among items for student pre-test, post-test, gains, and teacher ratings, respectively. Let $L^2{}_1$ be a diagonal matrix consisting of the first three largest latent roots of R_1, $L^2{}_2$, and $L^2{}_3$ be diagonal matrices consisted of the first two largest latent roots of R_2 and R_3, respectively, and finally $L^2{}_4$ be a diagonal matrix consisting of the first five largest latent roots of R_4. If we let A_1, A_2, A_3 and A_4 be the principal factor coefficients for each R, they are given by the equations, $A_1 = F_1L_1$, $A_2 = F_2L_2$, $A_3 = F_3L_3$, and $A_4 = F_4L_4$ and the results are shown in Tables 14 and 15.

As can be seen in the bottom row of these tables, the first latent roots for the teacher ratings and the student post-test performance are the largest of all the latent roots. This implies that the first principal factor of the teacher ratings accounts for the largest part of the variance in teacher ratings and also the first principal factor of the student post-test performance accounts for the largest part of the variance of the student post-test scores on items. Other factors are less dominant than these two factors. The magnitude of the latent root, then, indicates the degree of the importance of the factor.

The sum of the latent roots, however, is small for any of the student data with respect to the total variance. The sum of the three largest latent roots for the student pre-test is 10.704 and it accounts for only 19.8% of the total variance.* For the student post-test, the sum of the first two largest latent roots is 10.619 and it accounts for 19.7% of the total variance. For the student gain scores, the sum is 5.849, 10.8% of the total variance. The remainder of total variance seems to have no particular common factors. It is the sum of the uniquenesses for the items which include unreliability or error. These low communalities may result from the fact that each item correlated was composed of no more than four sub-items. For the teacher ratings, however, the sum of the first five largest latent roots considered as common factors is 33.195, and it accounts for 61.5% of the total variance of teacher ratings.

The similarities of the principal factor coefficients for the items at pre-test, post-test, and from the gain scores, with the principal factor coefficients of teacher ratings are given by the matrices $F_1'F_4$, $F_2'F_4$, and $F_3'F_4$ respectively. They are shown in Table 16, where each cell of a matrix includes an element of $F_1'F_4$, $F_2'F_4$, or $F_3'F_4$. The first principal factor of the teacher ratings has the largest similarity with the first factor of each of the three kinds of student data. The degree of similarity is largest for the post-test (.9281), second largest for the pre-test (.8785) and least for the gain scores (.5659).

---

* Since the variance of each item is standardized, the total variance is identical with the number of items, 54.

It seems interesting that the second factors for both pre-test and post-test scores have relatively large similarities with the fifth factor for the teacher ratings, although these similarities are not significant. For the gain scores, the second factor does not have a particularly large similarity with any factor for teacher ratings.

In order to determine the congruent space over student performance and teacher ratings, factor matchings were carried out for each of the three kinds of student scores with the teacher ratings by the method discussed in Chapter II. $G_{11} = (F_1'F_4)(F_1'F_4)'$, $G_{22} = (F_2'F_4)(F_2'F_4)'$, and $G_{33} = (F_3'F_4)(F_3'F_4)'$ are computed. Their latent roots $\Lambda^2{}_1$, $\Lambda^2{}_2$, $\Lambda^2{}_3$ and the associated unit latent vectors, V_1, V_2 and V_3 are determined. $G_{4(1)} = (F_1'F_4)'(F_1'F_4)$, $G_{4(2)} = (F_2'F_4)'(F_2'F_4)$, and $G_{4(3)} = (F_3'F_4)'(F_3'F_4)$ are also computed, and the equivalence of their positive distinct latent roots with $\Lambda^2{}_1$, $\Lambda^2{}_2$, and $\Lambda^2{}_3$, found in the preceding computations, are checked. The unit latent vectors $V_{4(1)}$ for $G_{4(1)}$, $V_{4(2)}$ for $G_{4(2)}$, and $V_{4(3)}$ for $G_{4(3)}$ are determined. These results are summarized in Tables 17 and 18.

The normalized factor coefficients of the matched factors for student performance on pre-tests and those for teacher ratings are given by computing $B_1 = F_1V_1$ and $B_{4(1)} = F_4V_{4(1)}$. $B_2 = F_2V_2$ and $B_{4(2)} = F_4V_{(2)}$ also give the normalized factor coefficients of the matched factors for student performance in post-tests and those for teacher ratings. $B_3 = F_3V_3$ and $B_{4(3)} = F_4V_{4(3)}$ are for student gains and for teacher ratings, respectively. Each of the matched factors will be denoted by lower-case Roman numerals i_1, ii_1, $ii_{4(1)}$, $ii_{4(1)}$, etc. Subscripts 1, 2, 3, and 4 denote the pre-test, post-test, gains and the teacher ratings, respectively. i_1 denotes, for example, the first matched factor for student performance on pre-tests with teacher ratings, and $i_{4(1)}$ denotes the first matched factor for teacher ratings with student performance on pre-tests and so forth. They are given in Tables 19, 20 and 21.

The values $H_{14(1)} = B_1'B_{4(1)}$, $H_{24(2)} = B_2'B_{4(2)}$ and $H_{34(3)} = B_3'B_{4(3)}$ give the coefficients of congruence for matched factors of the teacher ratings with the student performance in pre-test, post-test and the gains respectively, and they are shown at each bottom of Tables 19, 20 and 21. The coefficients of congruence $H_{14(1)}$, $H_{24(2)}$, and $H_{34(2)}$ are identical with the square roots of the latent roots $\Lambda^2{}_1$ for G_{11}, $\Lambda^2{}_2$ for G_{22}, and $\Lambda^2{}_3$ for G_{33}, respectively.

It is interesting that the transformation matrix V_2 in Table 18 is the identity matrix which means that the principal factors for the student performance in post-tests are themselves maximumly congruent with the teacher ratings. A similar tendency holds for the student pre-test case. The main diagonal elements of the transformation matrix V_i are approximately one, and the off-diagonal elements are approximately zero. Hence, the transformation matrix V_1 is approximately the identity matrix, in other words, the principal factors for the student pre-test are themselves almost maximumly congruent with the teacher ratings. This tendency is not strong for the student gains. In order to reach the congruent axes, the principal axes have to be rotated about +21° and the second axis has to be reflected. For the transformation matrices for teacher ratings, $V_{4(i)}$, $V_{4(2)}$, and $V_{4(3)}$, it should be noticed that the first principal factor is, in every case, the factor most closely related to the first matched factors. This is true because the direction cosines of the axes between I_4 and i_4 are greater than .95 and dominant in each case.

As seen in Tables 19 and 20, the first matched factor for the post-test has the highest coefficient of congruence with the first matched factor for the teacher ratings (.941), and the first matched factor for the pre-test has the second highest coefficient with the first matched factor for the teacher ratings (.930). The rest of the factors for student performance have smaller coefficients with the factors for teacher ratings. The first matched factor for gain scores has

.632 as the coefficient of congruence with the first factor for the teacher ratings, which would be considered significant but not highly so. In the first stage of this study, the highest congruence was hypothesized for student gains with teacher ratings, the second for post-test and the last for pre-test. This hypothesis was not verified for student gains with teacher ratings, although the coefficient of congruence was higher for the post-test than for the pre-test as was hypothesized.

The minimum value of the coefficient of congruence such that two factors are regarded as meaningfully congruent may be chosen, more or less, on an arbitrary basis as in the case of setting the minimum correlation coefficient by which two variables would be regarded as highly correlated. If .60 is taken as the minimum coefficient, in such a sense, the second factors for student scores and teacher ratings will not be regarded as congruent in all the three cases. If .50 is taken as the minimum meaningful coefficient, the second factor for the post-test with the teacher ratings will be regarded as meaningfully congruent. If .40 is taken, the second factors for the student gains and the teacher ratings are regarded as congruent for all the three cases, but this criterion seems too small for a coefficient of meaningfully congruent factors. The variance accounted for by a coefficient of this size is only .16.

In order to see the relationships between matched factor coefficients visually, the pairs of the first normalized congruent factor coefficients are connected with lines as shown in Figures 6A, 6B, and 6C. In other words, the pairs of the values of the columns $i_{4(1)}$ and i_1 in Table 19 are plotted on the vertically parallel lines, and the points for each pair of values are connected as in Figure 6A. The same type of graph was made for the pairs of the values of the columns $i_{4(2)}$ and i_2 in Table 20, and for those of $i_{4(3)}$ and i_3 in Table 21, which are given in Figure 6B and 6C, respectively. If the two factor coefficients matched are identical with each other, all the lines connecting them should be horizontal and parallel. For the case of the post-test, which has the highest coefficient of congruence (.941), the lines lie most nearly together and horizontal, and the factor coefficients for the items are less scattered (Figure 6B). For the case of the pre-test which has the second highest coefficient of congruence (.930), the lines are more diverging. In particular, the factor coefficients for the student performance diverge more than those for the teacher ratings. This tendency is more evident for the case of the student gains with the teacher ratings in which the coefficient of congruence is .632. (See Figure 6C.) The factor coefficients for the gains scatter diversely while the factor coefficients for the teacher ratings do not.

For the matching of the second factors, the highest coefficient of congruence is seen for the second factor of the student post-test with the teacher ratings and it is .512. Figure 6D shows the factor relationships. However, a consistent pattern as seen in the cases of the first factors seems to have disappeared in this case. On this basis, the second matched factors are then considered to be non-congruent with each other for all three cases of student performance.* In the following discussion, therefore, only the first factors matched will be considered as candidates for the congruent factors. Also, as a consequence of the disappointing results for the gains scores correlations, no further consideration will be given to the factor analysis of these gains scores.

---

* The diagrams for the pre-test and the gain cases are not shown here but they look more random than in the post-test case.

TABLE 14

Principal Factor Coefficients of Student Tests*

Item Numbers	Pre-test Factors (A_1)			Post-test Factors (A_2)		Gain Factors (A_3)	
	I_1	II_1	III_1	I_2	II_2	I_3	II_3
Test I							
1	234	-188	227	351	-331	035	190
2	302	-138	069	244	-110	-246	-023
3	299	-312	-338	129	-270	-292	116
4	310	-032	-332	064	-087	-194	115
5	233	075	-075	201	-218	186	-043
6	-095	058	211	-057	243	-065	181
7	139	182	550	210	048	314	-037
8	141	263	463	308	095	414	-130
9	143	082	168	131	-044	168	025
10	165	-219	-010	060	-084	-133	139
11	110	-054	404	139	-078	-044	036
12	173	-386	189	434	-298	074	031
13	035	117	-070	303	-130	179	-158
14	051	-157	-095	445	-099	291	-235
15	033	-232	063	106	-094	-051	-268
16	134	-282	230	103	-172	-090	024
17	039	096	455	132	-056	315	017
18	-075	364	462	322	-104	500	-139
19	262	116	-164	278	-247	123	043
20	332	291	-213	350	-141	248	246
21	322	092	002	165	037	-001	273
22	343	459	-031	295	017	326	347
23	319	480	043	197	-199	268	133
24	399	324	-143	368	-220	256	398
25	142	209	130	240	-097	276	-006

---

*Decimal points are omitted.

TABLE 14 (continued)

Principal Factor Coefficients of Student Tests

Item Numbers	Pre-test Factors (A_1)			Post-test Factors (A_2)		Gain Factors (A_3)	
	I_1	II_1	III_1	I_2	II_2	I_3	II_3
Test II							
26	351	-034	-049	343	-060	-079	138
27	477	-094	-084	477	-108	086	129
28	193	-293	-007	442	-351	-147	344
29	254	137	175	513	-198	387	-285
30	423	-308	-211	504	-021	173	323
31	064	019	-212	475	-060	433	110
32	239	-231	-145	469	121	341	067
33	265	-266	236	544	-127	040	-165
34	-007	-046	187	449	-100	460	-191
35	048	-373	266	607	-183	-132	-189
36	407	-149	194	431	104	019	079
37	340	-341	356	505	-117	-045	-232
38	322	-335	122	460	-026	-111	035
39	423	-010	-005	365	-089	166	097
40	307	-299	-059	514	-083	180	056
41	-164	-063	-050	175	341	-043	-004
42	453	-237	144	575	148	008	114
43	462	-204	134	609	014	289	135
44	329	013	-052	530	082	227	133
45	399	031	-262	550	131	151	213
46	456	048	-056	517	174	-000	270
47	481	203	-311	587	133	238	658
48	449	239	-300	365	439	138	612
49	307	-099	007	343	236	033	107
50	236	115	258	601	119	426	-129
51	392	412	081	363	591	399	-070
52	358	415	-142	512	353	573	-184
53	127	398	360	215	590	234	-466
54	411	347	039	222	550	107	-292
Latent Roots	4.860	3.144	2.700	8.041	2.578	3.230	2.619

TABLE 15

Principal Factor Coefficients of Teacher Ratings*

Item	Rating Factors (A_4)				
Numbers	I_4	II_4	III_4	IV_4	V_4
Test I					
1	409	-326	292	-013	-321
2	588	045	-070	-187	051
3	517	116	-222	306	415
4	580	216	-173	428	274
5	704	-259	-209	047	095
6	353	-301	248	250	446
7	700	171	-331	-036	271
8	665	166	-315	225	238
9	497	515	-024	-265	216
10	439	557	-087	-141	152
11	738	130	-144	-184	019
12	539	421	101	-114	-185
13	418	670	-213	-140	-195
14	417	643	-031	112	-080
15	549	137	070	318	-246
16	602	281	240	092	-282
17	408	700	-056	-136	-090
18	340	787	026	-041	-062
19	603	217	528	-089	-043
20	539	215	625	080	-124
21	402	144	705	048	-050
22	399	298	544	015	111
23	569	122	421	032	-024
24	714	-093	-113	207	168
25	570	-321	460	084	112

---

*Decimal points are omitted

TABLE 15 (Continued)

Principal Factor Coefficients of Teacher Ratings

Item Numbers	Rating Factors (A_4)				
	I_4	II_4	III_4	IV_4	V_4
Test II					
26	635	-185	-060	007	-310
27	669	-190	-009	246	-270
28	539	-331	252	231	-094
29	713	-135	-076	401	-228
30	756	-067	047	242	-129
31	735	-080	-374	306	-115
32	626	044	-121	123	-293
33	697	-011	-268	349	-128
34	659	261	-076	271	007
35	726	-312	-052	-272	-178
36	740	-248	-129	-229	-139
37	735	-217	-056	-287	-127
38	701	-002	168	-090	170
39	705	-144	023	-184	-040
40	802	-071	-014	-173	-096
41	722	-147	-353	-242	-109
42	769	-108	-305	-218	-050
43	756	-162	-163	-105	094
44	698	-288	-056	-268	055
45	691	-211	-023	-153	-184
46	614	-245	170	-204	-107
47	634	-071	093	-383	297
48	558	-285	263	-242	292
49	505	-242	303	153	159
50	569	-269	172	198	122
51	838	-088	-261	-013	072
52	798	-066	-164	011	166
53	486	230	360	-036	239
54	613	-047	061	-059	341
Latent Roots	20.770	4.594	3.564	2.241	2.026

TABLE 16

Principal-Factor Similarities between Student Performance and Teacher Ratings

Student Performance	I_4^*	II_4	III_4	IV_4	V_4
Pre-test (F_1')					
I_1	.8785	-.1738	.1554	-.0985	.0328
II_1	-.0234	.1423	.2359	.0002	.3900
III_1	.2089	.2093	-.0650	-.0980	-.0379
Post-test (F_2')					
I_2	.9281	-.0801	.0347	-.0409	.1191
II_2	.0295	-.1860	.0605	-.2292	.4135
Gains (F_3')					
I_3	.5659	.1971	.0208	.0978	.0447
II_3	.2097	-.2266	.2980	-.1482	.1180

* Roman numerals indicate the principal factors. Arabic subscripts indicate the kind of data from which the factors were obtained.

TABLE 17

Cross-Products of the Principal-Factor Similarities

Pre-test over
Teacher Ratings
$G_{11} = (F_1'F_4)(F_1'F_4)'$

	I_1	II_1	III_1
I_1	.8370	.0043	.1454
II_1	.0043	.2285	-.0053
III_1	.1454	-.0053	.1027

Teacher Ratings
over Pre-test
$G_{4(1)} = (F_1'F_4)'(F_1'F_4)$

	I_4	II_4	III_4	IV_4	V_4
I_4	.8160	-.1123	.1174	-.1070	.0119
II_4	-.1123	.0942	-.0071	-.0033	.0418
III_4	.1174	-.0071	.0840	-.0089	.0996
IV_4	-.1070	-.0033	-.0089	.0193	.0005
V_4	.0119	.0418	.0996	.0005	.1546

Post-test over
Teacher Ratings
$G_{22} = (F_2'F_4)(F_2'F_4)'$

	I_2	II_2
I_2	.8849	.0004
II_2	.0004	.2626

Teacher Ratings
over Post-test
$G_{4(2)} = (F_2'F_4)'(F_2'F_4)$

	I_4	II_4	III_4	IV_4	V_4
I_4	.8623	-.0799	.0303	-.0448	-.0983
II_4	-.0799	.0410	.0085	.0459	-.0674
III_4	.0303	.0085	.0049	.0124	-.0291
IV_4	-.0448	.0459	.0124	.0542	-.0899
V_4	-.0983	-.0674	-.0291	-.0899	.1851

Gains over
Teacher Ratings
$G_{33} = (F_3'F_4)(F_3'F_4)$

	I_3	II_3
I_3	.3711	.0710
II_3	.0710	.2201

Teacher Ratings
over Gains
$G_{4(3)} = (F_3'F_4)'(F_3'F_4)$

	I_4	II_4	III_4	IV_4	V_4
I_4	.3643	.0639	.0742	.0243	.0501
II_4	.0639	.0902	-.0635	.0529	-.0180
III_4	.0742	-.0635	.0892	-.0421	.0361
IV_4	.0243	.0529	-.0421	.0315	-.0131
V_4	.0501	-.0180	.0361	-.0131	.0159

TABLE 18

Transformation Matrices and Latent Roots

Pre-test V_1

	i_1	ii_1	iii_1
I_1	.9823	-.0023	-.1875
II_1	.0050	.9992	.0387
III_1	.1874	-.0390	.9815

Post-test V_2

	i_2	ii_2
I_2	1.0000	-.0006
II_2	.0006	1.0000

Gains V_3

	i_3	ii_3
I_3	.9296	.3685
II_3	.3685	-.9296

Teacher Ratings $V_{4(1)}$

	$i_{4(1)}$	$ii_{4(1)}$	$iii_{4(1)}$
I_4	.9700	-.0614	.1442
II_4	-.1407	.2793	.8909
III_4	.1522	.4988	-.3068
IV_4	-.1238	.0078	-.2839
V_4	.0292	.8181	-.1036

Teacher Ratings $V_{4(2)}$

	$i_{4(2)}$	$ii_{4(2)}$
I_4	.9867	.0565
II_4	-.0853	-.3629
III_4	.0368	-.1181
IV_4	-.0437	-.4472
V_4	-.1263	.8069

Teacher Ratings $V_{4(3)}$

	$i_{4(3)}$	$ii_{4(3)}$
I_4	.9550	.0309
II_4	.1576	.6467
III_4	.2042	-.6149
IV_4	.0575	.3967
V_4	.1347	-.2130

Latent Roots of Matrix G_{11}

Λ_1^2

.8648	.0000	.0000
.0000	.2287	.0000
.0000	.0000	.0747

Latent roots of Matrix G_{22}

Λ_2^2

.8849	.0000
.0000	.2626

Latent roots of Matrix G_{33}

Λ_3^2

3993	.0000
.0000	.1919

TABLE 19

Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Pre-test Performance*

	Teacher Ratings $B_{4(1)} = F_4 V_{4(1)}$			Student Pre-test $B_1 = F_1 V_1$		
Items	$i_{4(1)}$	$ii_{4(1)}$	$iii_{4(1)}$	i_1	ii_1	iii_1
Test I						
1	126	-156	-144	130	-111	112
2	133	008	081	142	-079	013
3	068	190	012	094	-168	-234
4	065	134	035	100	-010	-225
5	148	-043	-067	095	045	-063
6	103	279	-234	-018	028	136
7	120	081	134	125	090	320
8	092	067	081	116	137	270
9	096	177	268	083	042	090
10	065	130	276	072	-123	-025
11	153	-022	134	095	-040	231
12	101	-032	211	097	-222	090
13	036	-088	367	008	068	-042
14	033	025	270	011	-086	-064
15	082	-111	020	021	-133	030
16	116	-070	100	085	-164	120
17	046	018	345	070	043	271
18	025	069	346	020	194	290
19	163	135	043	098	070	-118
20	142	115	-001	124	169	-149
21	128	171	-048	144	052	-024
22	110	241	037	151	260	-038
23	144	106	-004	149	270	009
24	135	046	-049	162	186	-112
25	132	140	-221	079	115	070

---

* Decimal points are omitted.

Coefficients of Congruence between Above Factors

$H_{14(1)} = B_1' B_{4(1)} = \Lambda_1$

	$i_{14(1)}$	$ii_{14(1)}$	$iii_{14(1)}$
i_1	.9299	.0000	.0000
ii_1	.0000	.4782	.0000
iii_1	.0000	.0000	.2733

TABLE 19 (continued)

Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Pre-test Performance*

Items	Teacher Ratings $B_{4(1)} = F_4 V_{4(1)}$			Student Pre-test $B_1 = F_1 V_1$		
	$i_{4(1)}$	$ii_{4(1)}$	$iii_{4(1)}$	i_1	ii_1	iii_1
Test II						
26	135	-227	-026	151	-018	-060
27	128	-190	-083	203	-050	-093
28	136	-037	-199	085	-165	-027
29	117	-176	-081	133	073	086
30	146	-079	-048	163	-168	-169
31	104	-184	001	005	016	-132
32	104	-203	056	089	-126	-112
33	096	-154	004	144	-155	112
34	095	011	090	018	-030	112
35	190	-168	-034	244	-216	110
36	179	-158	-005	203	-088	078
37	187	-127	006	191	-200	176
38	174	132	-001	156	-192	038
39	176	-046	-004	188	-005	-039
40	186	-080	038	129	-167	-068
41	153	-186	073	-079	-034	-017
42	163	-135	074	218	-136	042
43	169	-021	-004	220	-118	036
44	186	-032	-042	141	009	-059
45	168	-149	-020	148	024	-190
46	175	-058	-064	197	029	-071
47	185	176	027	180	122	-222
48	185	192	-119	167	142	-212
49	138	134	-174	137	-056	-024
50	139	074	-168	135	059	137
51	166	-051	030	185	231	024
52	163	032	010	145	238	-106
53	125	256	042	099	216	213
54	150	197	-024	189	195	-004

* Decimal points are omitted.

TABLE 20

Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Post-test Performance*

Items	Teacher Ratings $B_{4}(2) = F_{4}V_{4(2)}$ $i_{4(2)}$	Teacher Ratings $B_{4}(2) = F_{4}V_{4(2)}$ $ii_{4(2)}$	Student Post-test $B_2 = F_2V_2$ i_2	Student Post-test $B_2 = F_2V_2$ ii_2
Test I				
1	136	-136	124	-206
2	125	088	086	-069
3	057	144	045	-168
4	077	009	022	-054
5	149	106	071	-135
6	046	218	-020	151
7	115	165	074	030
8	104	067	109	059
9	075	122	046	-028
10	062	045	021	-052
11	155	062	049	-049
12	122	-142	153	-186
13	081	-164	107	-081
14	068	-180	157	-062
15	127	-255	037	-059
16	146	243	036	-107
17	071	-120	046	-035
18	050	-154	113	-065
19	139	-060	098	-154
20	129	-163	123	-088
21	098	-106	058	023
22	075	-212	104	010
23	128	-063	069	-124
24	135	065	130	-137
25	089	068	085	-061

Items	Teacher Ratings $B_{4(2)} = F_2V_2$ $i_{4(2)}$	Teacher Ratings $B_{4(2)} = F_2V_2$ $ii_{4(2)}$	Student Post-test $B_2 = F_2V_2$ i_2	Student Post-test $B_2 = F_2V_2$ ii_2
Test II				
26	171	-135	121	-037
27	169	-185	168	-067
28	136	-075	156	-219
29	167	-213	181	-123
30	172	-128	178	-013
31	156	-111	167	-038
32	154	-195	165	075
33	147	-150	192	-079
34	122	-108	158	-062
35	192	046	214	-114
36	186	049	152	065
37	186	063	178	-073
38	143	121	162	-016
39	168	064	129	-056
40	190	020	181	-052
41	172	066	062	213
42	[illegible]	084	203	092
43	[illegible]	[illegible]	215	009
44	164	[illegible]	187	051
45	178	-125	194	081
46	161	039	182	108
47	127	297	207	083
48	118	277	129	274
49	106	073	121	147
50	121	052	212	074
51	174	086	128	368
52	157	122	181	220
53	083	091	076	367
54	107	222	078	342

Coefficients of Congruence between Above Factors

$H_{24(2)} = B'_2 B_{4(2)} = \Lambda_2$

	$i_{4(2)}$	$ii_{4(2)}$
i_2	.9407	.0000
ii_2	.0000	.5124

---

* Decimal points are omitted.

TABLE 21

Normalized Congruent-Factor Coefficients for Teacher Ratings and Student Gains*

Items	Teacher Ratings $B_{4(3)} = F_4 V_{4(3)}$		Student Gains $B_3 = F_3 V_3$	
	$i_{4(3)}$	$ii_{4(3)}$	i_3	ii_3
Test I				
1	063	-146	062	-102
2	117	-017	-132	-037
3	144	130	-125	-126
4	161	198	-074	-106
5	117	-007	086	063
6	131	-170	008	-118
7	148	114	154	086
8	149	181	184	160
9	150	064	093	020
10	132	139	-037	-107
11	143	040	-014	-030
12	133	095	045	-003
13	090	267	057	128
14	129	249	097	195
15	122	143	-088	144
16	150	077	-041	-032
17	117	210	167	055
18	125	230	227	182
19	192	-120	073	001
20	188	-095	185	-091
21	168	-163	062	-157
22	176	-097	247	-132
23	173	-084	169	-021
24	154	043	223	-177
25	117	-239	141	060

Items	Teacher Ratings $B_{4(3)} = F_4 V_{4(3)}$		Student Gains $B_3 = F_3 V_3$	
	$i_{4(3)}$	$ii_{4(3)}$	i_3	ii_3
Test II				
26	084	016	-009	-095
27	109	05[illegible]	074	-056
28	116	-103	002	-228
29	125	130	135	243
30	156	053	163	-150
31	108	201	249	025
32	098	133	192	031
33	118	200	-017	103
34	160	179	194	204
35	096	-118	-111	082
36	101	-068	028	-041
37	109	-099	-076	124
38	178	-100	-050	-043
39	129	-089	108	-022
40	146	-043	106	005
41	083	028	-024	-006
42	107	022	030	-064
43	134	-033	180	-018
44	114	-143	148	-030
45	103	-065	127	-091
46	111	-163	061	-155
47	151	-193	273	-329
48	143	-276	234	-314
49	142	-151	041	-055
50	137	-099	191	161
51	147	050	190	121
52	161	017	255	223
53	179	-090	015	316
54	161	-096	-011	190

Coefficients of Congruence between Above Factors

$H_{34(3)} = B_3' B_{4(3)} = \Lambda_3$

	$i_{4(3)}$	$ii_{4(3)}$
i_3	.6319	.0000
ii_3	.0000	.4381

---

* Decimal points are omitted.

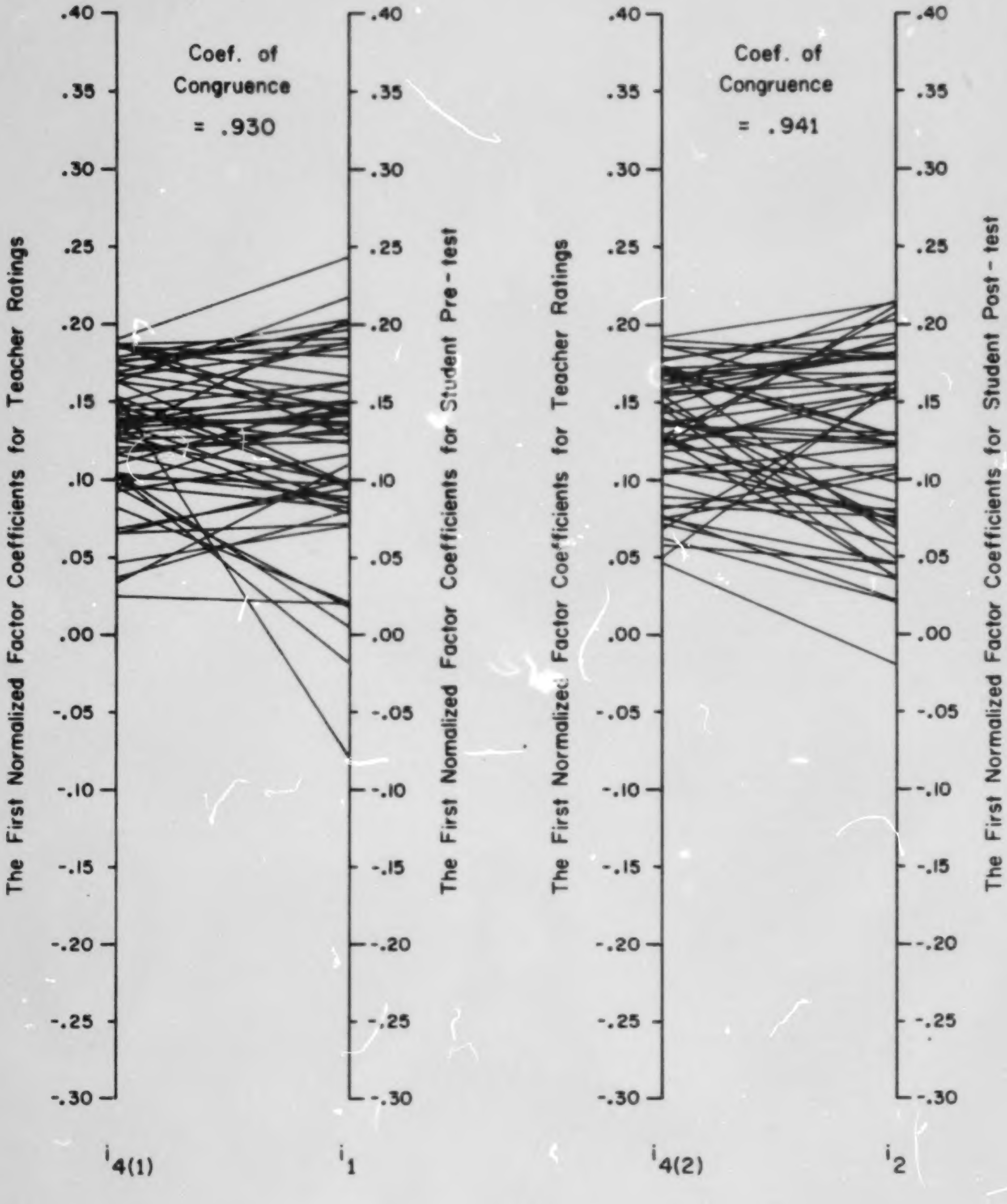


Figure 6

Matching of the Normalized Congruent-Factor Coefficients

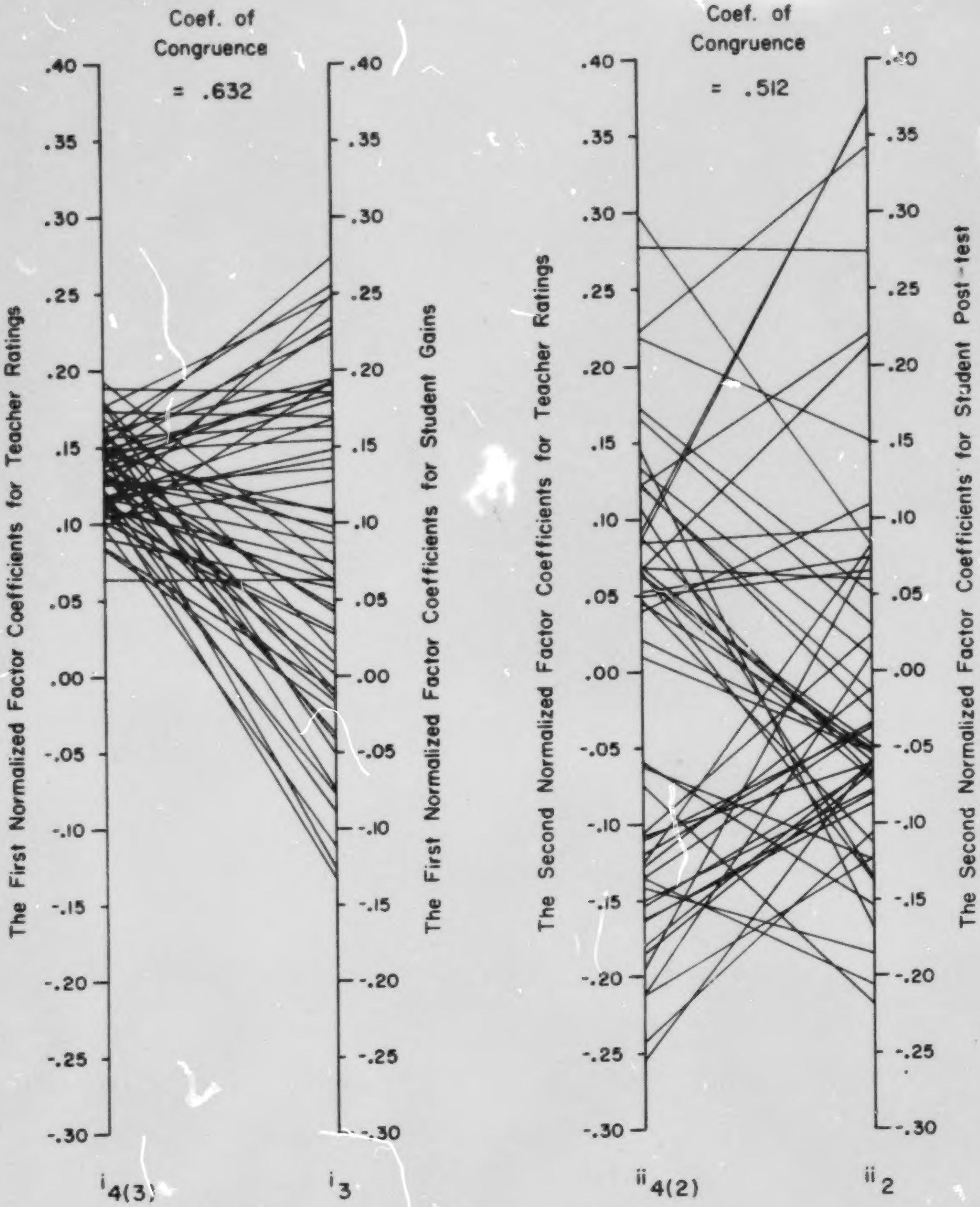


Figure 6 (Continued)

Matching of the Normalized Congruent-Factor Coefficients

CHAPTER V

DISCUSSION

1. Interpretation of the Results

Mean Student Performance

Looking at the mean student performance at the pre-test administrations, when instruction had just started, the students correctly answered more of the sub-items than we expected (45% of all the sub-items). Since at the post-test administration, the students correctly answered 67% of all the sub-items, it does not necessarily mean that the items prepared were too easy for the sample of students, but it rather means that the gains through the course were not great. In Figure 1, any items which would appear in the upper left corner would be difficult at the beginning of the instruction but easy by the time instruction was completed so that they would show large gains. However, there were only two such items, Items 7 and 8. Most of the items in the tests appear closer to the main diagonal which indicates no gain.

The items which were too easy at the beginning of the instruction allow no room for gains after instruction. Items 2, 3, 9, 11, 19, 26, 36 and 41 may be of this kind. However, the ceiling effect is clearly evident only for Items 11 and 41. Items 2, 3, 26 and 36 might have had more gains without the ceiling, but they show fairly good gains, considering other items, so the ceiling effect may not have excessively inhibited the potential gains on these items. In spite of having more space for gains, Items 9 and 19 did not show much. Other items having statistically non-significant gains at the 5% level such as Items 6, 10, 13 and 20 were correctly answered at about 50% for both pre-test and post-test administrations. Item 34 was difficult for both pre- and post-test administrations, and it had no great gain in spite of ample possibility. Thus, the low gains in general evidently are not because the test items were easy but because the improvement by instruction was not great.

It should be noticed, however, that many items whose contents were not taught between the pre-and post-test administrations were included for the experimental purpose. Low gains, then, do not necessarily mean the ineffectiveness of the teachers' instruction or training to their students. Five items out of seven whose gains are not significantly different from zero are classified as least related to the test. (See Table 3.) This suggests that the degree of relevance should have some relation with mean gain scores. Following the classification in Table 3, the means of mean gain scores for items were computed for each group of items (See Table 22). The mean is 1.02 for Group A, in which the contents of items are most closely related to the text. The mean is .98 for Group B, in which the contents of items are moderately related to the text. The two means for Groups A and B are not greatly different but the mean is .54 for Group C, in which the contents of items are least related to the text. This mean gain score is quite low compared with other two groups of items. The contents of the items which belong to the first two groups are related, to some extent, to what is taught between the two administrations. But for the most of the items of Group C, the contents were not taught between the two tests. Therefore, the low mean of the gains for Group C agrees with what we would expect from common experience.

Looking at the contents of Items 7, 8, 12, 17, 27, 40 and 47, which have larger gains than the others, four items out of seven are those which require students to have some knowledge of computational rules. The rest of the items, 12, 40 and 47, also seem to require some drill after learning the relevant mathematical principles. Items which can be answered without the effort of thinking about principles do not seem to be included in this group. In other words, these items would not be answered by simple rote learning of mathematical concepts such as " 'APA' is an abbreviation of the Associative Principle for Addition" or "The equation '$^{+}2 + {}^{-}3 = {}^{-}3 + {}^{+}2$' is an instance of the Commutative Principle for Addition," for instance.

Following the classification of Table 2, the means of the mean gain scores for each group are computed (See Table 22). The mean is .72 for Group A in which the questions are asking for direct understandings of basic concepts. The mean is 1.09 for Group B in which the questions are asking some computational work, and the mean is .92 for Group C in which some applied work using basic concepts is required. The means of the gains are the highest for Group B, the second for Group C and the least for Group A. This finding also supports the conclusion that more progress was made during the course on items requiring some drill-like computation or application of learned principles than on those requiring simple knowledge of mathematical terms or simple recognition of instances of mathematical concepts.

Also, if we compute the means of the mean gain scores for each group of items, classified by format according to Table 1, the mean for the multiple-choice type is .76, the mean for the numerical completion type is 1.13 and that for the type requiring written work is .88. The highest gain is obtained for the group of items which can be answered by filling with numerals or algebraic variables, and the lowest gain is obtained for the group of multiple-choice type items. This classification is not independent, however, of the content of the items. As we have seen in the preceding paragraph, the items asking for simple basic concepts had the lowest gains, and more than two thirds of such items were multiple-choice type. The items asking for some computational work had the highest gains, and also more than two thirds of such kinds were

TABLE 22

Mean Student Gains and Teacher Ratings for Different Types of Items

Answer Forms	Number of Items	Mean Student Gains	Mean Teacher Ratings
A: Multiple-choice	26	.76	5.94
B: Filling with Numerals or Algebraic Letters	17	1.13	5.64
C: Written Work	11	.88	5.82
Required Abilities			
A: Understanding of Basic Concepts	19	.72	5.80
B: Computational Skill	16	1.09	5.63
C: Ability of Application of Basic Concepts	19	.92	6.00
Degree of Relevance to the Text			
A: Closely Related	22	1.02	5.79
B: Moderately Related	20	.98	6.11
C: Least Related	12	.54	5.39
Grand Means over All the Items	54	.90	5.82
Standard Deviation of Means over All the Items	54	.63	.49

simply answerable by writing numerals representing the outcomes of the computational work. Therefore, the differences of the mean gains would be the consequences of the nature of the tasks rather than of the answer forms.

To summarize the student mean performances, a large number of students correctly answered the test items before the contents were taught. However, the progress in the course was not as great as was expected. It is partly because some of the test contents were not taught between the pre-and post-test administrations. The highest gain was obtained in the tasks which require some drill or exercises on computation or applications of the basic principles studied in the course. In the tasks which require simple knowledge of mathematical terms or recognition of mathematical concepts, the gain was low. For the tasks whose contents were not taught between the pre- and post-test administrations, the gain was the lowest.

Mean Teacher Ratings

The next question concerns how teachers evaluate the test items. As can be seen in Table 12, the items which teachers like most are items 24, 35, 37 and 45. Judging from these items, it would seem that teachers like items that (1) are not directly drawn from the text but whose underlying concepts are closely related to it (Items 24 and 45), (2) have a form and content similar to problems in the text but whose instances are new (Items 35 and 37), and (3) include some computational work in which computation itself is not a main goal but only a step to reach other mathematical concepts (Items 24, 35, 37 and 45). In short, teachers seem to like challenging problems. But this tendency is not supported by all the data.

Looking at the items rated lowest, such as Items 13, 14, 17, 18, 21 and 22, it seems obvious that teachers tend not to like items whose contents are not taught in the course, and this is quite reasonable. However, disregarding such obviously inadequate items, teachers seem not to like items that have the following properties: (1) simple computational work (Item 3), (2) straightforward and directly related to the text (Items 2 and 41), (3) improper wording (Item 22*), (4) too wordy for the expected outcomes (Item 53), and (5) containing metamathematics rather than conventional mathematics (Item 1). These tendencies also are not fully supported by the evidence, since there are many items which are not rated low that have the above properties.**

If we compute the means of the mean teacher ratings for each type of item as we did for student gains, the last column of Table 22 is obtained. The teachers do not like the items whose contents are least related or irrelevant to the text (5.39). They like the items whose contents are somewhat modified from the examples in the text (6.11) more than questions taken directly from the text (5.79). The teachers also seem to like questions which require some applied work using basic concepts in the text (6.00) and they like less those requiring simple computation (5.63) or direct questions on the basic mathematical concepts (5.80). The teachers seem to like multiple-choice items (5.94) more than simple writing of numerical values or algebraic variables (5.64). It may be partly because a number of the computations which are disliked by teachers are included in the latter category. The difference between the multiple-choice type (5.94) and the written-work type (5.82) is not great.

These general tendencies inferred from the ratings on the test items would be strengthened by investigating the teachers' responses to the questions on their teaching objectives and preferences on test construction.

As seen in the questionnaire (Appendix B), Item 9 in the preliminary part of the questionnaire asks the teachers to rank several objectives according to their importance, in their opinion. The average of the ranks for each objective are given in Table 23. A ranking for the entire sample was obtained by ordering the average ranks over all teachers. The teachers think that such abilities as understanding mathematical concepts, discovering mathematical relationships, and deductive reasoning are more important than skill in numerical computation and remembering mathematical principles. These findings are consonant with the earlier finding that the teachers tend to evaluate more highly mathematically sound items than those simply requiring numerical work or rote memory of mathematical concepts. The teachers rank last the ability to apply mathematical skills to real life problems. This must be the final goal of the mathematics education, but most teachers seem to think it a goal remote from everyday teaching. Kendall's coefficient of concordance (Kendall, 1955) may be a good index to show how the teachers' rankings resemble each other, and a value of W = .52 is obtained which appears to be a fairly high degree of concordance.

---

* The term 'arithmetic value' is used in the text instead of 'absolute value'.

** With some exceptions, the general tendencies described also hold for the four Pekin teachers. (See Table 12.)

Table 24 shows the mean values and the standard deviations of the teachers' preferences on test construction based on Item 10 of the questionnaire. The task given was to mark the proportion of items according to the key words shown in the table. If the mean proportion is less than 50.0, it means that the teachers prefer items characterized by the right-hand word to those characterized by the left-hand word. If the mean value is greater than 50.0, the relation is the reverse. This table should then describe the teachers' ideas or expectations concerning test construction. The teachers prefer relatively difficult items to relatively easy ones. They prefer the multiple-choice type to the completion type. They like non-verbal items and items with few words better than items with many words. They like quick response straightforward, and familiar questions better than time-consuming, tricky, and unfamiliar questions. They

TABLE 23

Ranks of Abilities Stressed by the UICSM Teachers

Total Sample Ranking	Average Ranks	Abilities
1	1.57	Understanding mathematical concepts
2	2.59	Discovering mathematical relationships
3	4.05	Deductive reasoning
4	5.05	Skill in symbolic manipulation
5	5.08	Generalizing from concrete objects to abstract ideas
6	5.27	Skill in numerical computation
7	5.62	Remembering mathematical principles
8	6.78	Applying mathematical skills to real life problems

TABLE 24

The Preferred Proportions of Items of Different Kinds to be Included in a Test for Grading

Left Key Words	Means of Percentages	Standard Deviations of Percentages	Right Key Words
relatively easy	48.1	16.1	relatively difficult
multiple-choice	54.1	19.5	completion
verbal	43.8	19.5	non-verbal
many words	40.9	23.2	few words
time-consuming	42.5	21.6	quick response
straightforward	66.9	25.8	tricky
familiar	66.0	22.0	unfamiliar
basic	60.5	16.4	applied
concrete	58.0	17.7	abstract
computational	51.8	16.5	conceptual
reasoning	60.4	18.4	recall
inductive	49.2	15.4	deductive
teacher constructed	50.2	23.6	UICSM project constructed

like basic, concrete, computational, reasoning, and deductive type questions better than applied, abstract, conceptual, recall and inductive type questions. Finally, they want to include teacher-made items and UICSM-made items in a half-and-half ratio. Of course, the degree of preference is different for each question, and it may be noted by the deviation of the mean value from 50.0. Of course, the mean value itself does not tell whether the preferences of the teachers are similar or diverse, but the standard deviation of the teachers' preferences would indicate this fact.

In most cases, the teachers' preferences agree with (or at least do not contradict) the findings from the teachers' ratings on the individual test items, although some caution is necessary. It would be understandable that they like reasoning, deductive, and familiar items better than recall, inductive,* and unfamiliar items, but the fact that they like basic, computational, and non-verbal items better than applied, conceptual, and verbal items appears contradictory to what we have reported in the earlier part of the study. The teachers respond to the questions here by comparing the given bipolar words, and their responses are a function of the given words. For example, the 'basic vs. applied' of which the teachers are thinking here must be different from the 'basic' items and the 'applied' items that we have classified in the earlier study. In the former case, teachers must have interpreted 'basic' as meaning a sound understanding, and a pre-requisite for applied work. In the latter case, however, by 'basic' was meant questions taken directly from the text, and 'applied' meant an ability to work new types, which is the test of a sound understanding of the underlying concepts. A similar argument may hold for the question 'straightforward' vs.

---

* The UICSM authors emphasize both deductive and inductive reasoning.

'tricky'. However, comparing the stimulus words on both sides, most of the teachers' responses seem understandable.

The next problem is to see the relationship between mean teacher ratings and mean student performances for each item. Before investigating the relationships between two kinds of responses from different subject groups, we should bear in mind that one assumption was involved. The samples of teachers were from 70 different schools in 19 states, representing a wide area in the United States. On the other hand, the sample of the students consisted of six classes, taught by four teachers, from one local high school in Illinois. When we are talking about teacher-student relationships in terms of responses on test items, we are interested in only the common tendencies that appeared in both teacher and student groups. We are interested in whether most students have or have not commonly shown gains, for instance, on items which most teachers commonly emphasize. We do not assume that any particular emphases different from other classes were made in these sample classes. Of course, if the teachers who are teaching the classes which are used for this study have excessively different objectives and different ways of teaching from other teachers, the teacher-student relationships which we are going to discuss here would loose its validity. However, since there is no reason to believe that these teachers had biased objectives from others, these teachers were assumed to be typical teachers in the United States. The years of teaching experience of mathematics for these teachers are four, five, twelve, sixteen years, and the years of teaching Units I and II of UICSM old texts are one, two, seven, and eight years, respectively. They took three, two, three, and two courses, respectively, in which the contents of UICSM curriculum were studied. Ideally, when we talk about the relationship between teacher-held objectives and student performance, the student performance should refer to the performance in the class taught by the teacher under consideration. We should find the relationships between the teacher-held objectives and his student performance. However, it would require a much longer and larger-scale research for this purpose.

Now, as seen in Figures 2 and 3, no particular tendencies are found between mean teacher ratings and mean student performances in the pre- and the post-tests. But a more important question is to ask if the students develop their ability along the line that the teachers think important.

To some extent, the answer is yes, but not completely so. As we have seen in the first section of this chapter, the students progressed more in applied work than in basic conceptual work, and the teachers also thought the former would be better than the latter in the ratings for items. The students showed more progress for the modified questions than for questions that were the same as in the text. The teachers also liked modified questions better than direct ones. The correlation of the mean student gain scores with the teacher ratings over items was .53 if a few exceptional items were taken out (Figure 4). This is not a strong relationship but there is a positive tendency between the teachers' objectives and the students' progress.

There is also a discrepancy between them. The teachers think that simple computational work is not so important compared with other aspects, as far as the items presented are concerned, but the students showed the greatest progress in this kind of work. Since the computational work requires some knowledge of operational rules and drill, the progress due to instruction would be most effective. Without knowledge of the rules and drill we could not expect the student to progress in mathematics, and this effect seems most evident in such a task as computation. The teachers replied to Item 10 of the questionnaire, indicating that they would include computational items more than 50% in their test, on the average. The teachers must recognize, in general, the importance of such drills as a basis for developing the students' ability in mathematics.

To summarize the teacher ratings of the items, the teachers emphasize, in general, the ability of basic and conceptual understanding of mathematics and its application rather than the rote memory or mechanical application of mathematical rules. They like questions modified from the examples of the text, but the questions should be based in the principles taught. They would exclude the questions whose solutions are based on principles not taught. This is reasonable, and in fact the students showed little progress in such questions. The teachers' objectives inferred from the mean ratings for the items are, in many cases, understandable in the light of their responses to the questionnaire asking for their general teaching objectives and attitudes toward test construction, although some caution is necessary. There is no particular relationship between the teacher ratings and the student performances on the pre- and the post-test, but there are positive but not strong tendencies between the mean ratings and the mean gain scores. The student progress tends to be large in the items which the teachers think important, although there are some exceptions, as seen in the computational type items.

Principal Factors for Student Performance

As mentioned in the previous chapter, we have obtained three factors from the intercorrelations among items for student pre-test scores and also two factors from the student post-test scores. For the gain score the factor analysis results were more ambiguous than for the pre-test scores and post-test scores, the first two principal factors accounted for only 10.8% of the score variances. Also, since the factor analysis of the student gains failed to yield interesting results, it will not be discussed.

The principal axes could be rotated by some criterion like the varimax (Harman, 1960). However, the interpretation of the factors will be done without rotation, since the principal axes themselves, particularly the first principal axis, have several significant characteristics as will be seen later. Varimax rotations were tried and did not lead marked improvements in the interpretability of the factor space.

For the pre-test case, the first factor refers to the ability which students already had when instruction of the text started. For the post-test case, the first factor implies the students' general ability on the given tests whose contents have been taught.

Judging from the contents of the items, the first principal factor for the pre-test seems to be related to a verbal vs. numerical dimension. The majority of items with low coefficients are numerical or computational type problems such as Items 13, 14, 17, 31 and 34. On the other hand, items with high coefficients on this factor seem to be related to the verbal description of mathematical problems or they are related to the ability of mathematical formulation from verbal statements, as seen in Items 24, 30, 42, 43, 46, 47, 48 and 54. Another important characteristic is that the items with an illustrative example (which suggests, not how to write answers, but how to think) are likely to have high coefficients. For example, Items 22, 27, 30, 39, 42, 47 and 48.

The implication of these findings is important. As the first principal axis is generally close to the first centroid axis, the first principal factor coefficients indicate the degree of the relationship of each item with the mean of all the items. Suppose that these items are used for an aptitude test and the mean of the item scores over all the items is used for a measure of the student aptitude. The first principal factor coefficients indicate the degree of contribution of each item to the aptitude defined above. The items with low coefficients do not serve as good aptitude test items. From what we have found in the preceding paragraphs, we may conclude that the computational type items are not good for aptitude test items in this sense. Since computation requires some knowledge of rule for operation and some practice, it may not be an aptitude. On the other hand, the ability to formulate mathematical concepts described verbally in mathematical form may be a good indication of the possibility of reaching a high level of achievement in later learning. When some mathematical illustrations are presented, the ability to catch the mathematical essence from the illustrations and the ability to apply it to the given problems is also an aptitude for getting a high level of achievement in mathematics. Thus, the first principal factor coefficients of items for the pre-test seems to suggest which items are good ones for inclusion in an aptitude test.

The second principal factor for student pre-test seems to be a time-related effect. Items with high coefficients are located near the end of both Tests I and II (Items 22, 23, 24, 25, 51, 52, 53, and 54). Otherwise, items appear to need a lot of time to answer perhaps making students hesitate to do them first (Items 8, 18, 47 and 48). On the other hand, items with high negative coefficients are likely to be multiple-choice type which can be responded to quickly, or at least, these items look easy and most students are likely to attempt to do them first. Items 3, 12, 15, 16, 28, 35, 37, 38, 40, 42 and 43 are examples.

The third principal factor for the student pre-test is hard to interpret. Items with high coefficients are Items, 7, 8, 11, 17, 18, 37 and 53. They are problems involving multiplication, division, and the laws of operation such as the commutative laws, the associative laws, etc. Items with high negatively coefficients are Items 3, 4, 45, 47 and 48. They are problems in addition, mathematical formulation in using quantifiers, and finding the numbers of roots of equations. Since there is no consistent pattern, we would like to leave this factor unidentified.

For the post-test, the first principal factor seems to be closely related to the objectives of the given UICSM text. The items which have high coefficients on this factor seem to accord with what the text emphasizes in the given chapters.

Items 29, 30, 33, 35, 37, 40, 42, 43, 44, 45, 46, 47 and 50 are the examples. On the other hand, items with low coefficients, such as 9, 10, 15, 16, 17 and 21, are less emphasized in the given chapters. This is an important requirement for the achievement test, and the first principal factor coefficients seem to indicate which items are good ones for the achievement test. The behaviors of Items 6 and 41 are unique for both pre- and post-test data. Item 6 is a challenging type of problem and it does not follow the general pattern in many other cases. However, the low coefficient of Item 41 is difficult to understand.

The interpretation of the second principal factor for the student post-test is not easy. Items with high coefficients are Items 41, 48, 51, 52, 53 and 54, and items with high negative coefficients are 1, 3, 12, 19, and 28. Although no consistent patterns exist among them, Items 1 and 28 are rather inductive and straightforward type problems, while Items 51, 52, 53 and 54 appear to be logical reasoning and deductive-type problems. Looking at inductive-type problems such as Items 15, 25 and 39, they all have negative coefficients. However, as there are many deductive type items with negative coefficients, this 'inductive-deductive' interpretation is not reliable.

The analysis of student gain scores was not considered in this discussion due to the lack of interesting results.

The second factors are difficult to interpret for the student data, and this is not improved by rotation of axes. The reason for the difficulty of interpretation of the second factors is due to the fact that they are almost indistinguishable from random factors. The factors which we could use with some reliability are only the first ones. In fact, only the first factors are congruent with the factors for teacher ratings, when they are matched with teacher factors.

In summary, the first principal factor from student pre-test data is interpreted as a general aptitude factor in which items requiring an ability to grasp mathematical concepts from verbal statement have high coefficients and items requiring some knowledge of mathematical rules and practice have low coefficients on this factor. The second principal factor from student pre-tests is interpreted as a time factor in which time-consuming items or items located near the end of the tests have high positive coefficients, and items quickly responded to have high negative coefficients. The third principal factor for student pre-test is hard to interpret and it is left unidentified.

For the student post-test, the first principal factor is interpreted as a general achievement factor in which items stressed in the text have high coefficients on this factor. Items which are not treated in the chapters under consideration tend to have low coefficients on this factor. The second principal factor is interpreted as a deductive-inductive factor, but it is not a clear interpretation.

Principal Factors for Teacher Ratings

Five factors were extracted from the teacher ratings by the principal component method. The first principal factor is related to the teachers' individual differences on their ratings taken as a whole. As the first principal axis is expected to be close to the first centroid axis which represents the mean of the standard ratings of all the items, the first principal factor is highly related to the general tendency of each teacher's ratings. Therefore, it may be interpreted as reflecting the teachers' general response set for all the items. Items with high coefficients (greater than .70) on this factor are 5, 11, 24, 29, 30, 31, 35, 36, 37, 38, 39, 40, 41, 42, 43, 51 and 52. Most of them are items in Test II. These items (especially those in Test II) are mostly conceptual type items. Items which are irrelevant to the text are likely to have low coefficients on this factor. Some aspects of the relevance of items to the text are then confounded with the second and the third principal factors which will be described in the next paragraph.

Both the second and the third principal factors seem to be related to the relevance of items to the text, but the two factors are different in some characteristics. Items with high coefficients on the second factor are 9, 10, 13, 14, 17 and 18 and they are computational type problems on which the teaching is delayed to later chapters, but the questions themselves are quite straightforward. Items with high coefficients on the third factor are 19, 20, 21, 22, 23 and 25 whose contents are about inequalities, absolute values and symbolic manipulations. The contents of these items is taught only later or taught in a different way from what the items require. For example, the term "absolute value" in Item 22 is not used in the text; instead, the term "arithmetic value" is used. Items 21 and 25 are application problems of abstract symbols and the same kind of question is not practiced in the text.

Another important characteristics of the second factor is that this factor is related to conventional mathematics problems vs. new mathematics problems. Looking at items with high negative coefficients such as Items 1, 5, 6, 25, 28, 35, 36, 44, 46, 48, 49 and 50, they are dealing with new mathematic problems such as metamathematics, symbolic operations, generalization, open sentences, the principles of real numbers, quantifiers and so on. On the other hand, items with high positive coefficients are rather traditional skill-oriented problems. This characteristic is important but it overlaps with relevance to the text.

The fourth principal factor seems to refer to an algebraic factor. Items with high coefficient on this factor are likely to treat algebraic manipulation including algebraic variables, a, b, x, y, etc. Items 3, 4, 27, 28, 29, 30, 31, 33 and 34 are the examples. On the other hand, items with high negative coefficients are likely to be more conceptual as seen in Items 35, 36, 37, 41, 42, 44, 46, 47, and 48. They do not require actual algebraic operations.

The fifth principal factor is hard to interpret. Items with high positive coefficients are 3, 4, 6, 7, 8, 9, 47, 48, 53 and 54, while items with high negative coefficients are 1, 15, 16, 26, 27, 29 and 32. Comparing the contents of both groups of items, no systematic pattern is found. Hence, we had better leave this factor unidentified.*

Congruent Factors between Student Performance and Teacher Ratings

The next problem is to match the teacher factors with the student factors. As have seen in the earlier chapter, only one composite factor from student data is congruent with one composite factor from teacher data. Comparing two factor coefficients matched for student pre-test and teacher ratings, items with high coefficients for both student and teacher data are Items 35, 36, 37, 38, 39, 42, 43, 44, 46, 47, 48, 51, 52 and 54.** All are items in Test II and generally they are all closely related to the contents of the text. Items which have low coefficients for both teacher and student data are Items 3, 6, 9, 10, 12, 13, 14, 15, 16, 17, 18, 31, 32 and 34. Most of them are computational-type problems, and they require that students have some knowledge of rules for operation and practice. Items with great discrepancies between the two factors are Items 6 and 41. Both have high coefficients for teacher ratings and low coefficients for student pre-tests. Item 6 is a somewhat peculiar item, and its behavior is hard to understand. So is Item 41. Generally speaking, conceptual items and items closely attached to the text tend to have high coefficients on both teacher and student new factors. Computational type items which require students to have some knowledge and practice tend to have low coefficients on both teacher and student factors. Thus, this conceptual vs.

---

* If the factors were rotated, the interpretation might have been easier. However, rotation by the varimax criterion did not significantly increase the ease of the interpretations. The whole system of factor interpretation by the varimax method was much the same as that by the principal axes method.

** For easier comparison, the factor coefficients (see Table 19) are normalized over all the items.

computational factor is a congruent factor for teacher ratings and student performance at the pre-test stage. The coefficient of congruence is .930.

As for a congruent space over teacher ratings and student post-test data, it should be noticed that the student congruent factor is the same as the first principal factor. As seen in Table 18, the transformation matrix V_2 is the identity matrix and the result is not changed by the transformation from the first principal axis for post-test data to a new axis, maximally congruent with the corresponding new factor for teacher ratings. A similar argument holds for the transformation of the teacher matrix. The first row first column element of the matrix $V_{4(2)}$ in Table 18 is nearly one, and the rest of the elements in the first column are nearly zero. This means that the new teacher axis is not significantly changed from the original first principal factor axis for teacher ratings. In other words, the first principal factors for students and teachers are themselves almost maximumly congruent with each other. In fact, the coefficients of factor similarity between the first principal factors is .928 (Table 16), and it is close to the maximum coefficient of congruence .941 after transformation.

Items which have high coefficients on both data for students and teachers are Items 27, 29, 30, 31, 32, 35, 36, 37, 40, 42, 43, 44, 45, 46 and 52. They are closely related to the contents of the text. The difference from the pre-test case is that items which require some algebraic manipulation are now included in the post-test case. Items 27, 29, 30, 31, and 32 are the examples. To be accustomed to such algebraic manipulation may be necessary in order to answer these questions, and this would be expected as a consequence of study and practice based on the text. Items which have low coefficients on both kinds of data are 3, 4, 6, 9, 10, 13, 17, 21, 25 and 53. Most of them are the items whose contents are trivial or not taught between the pre-test and the post-test administrations. Thus the congruent factor for the teacher ratings and the student performance at the post-test time seems to be a factor which stands for a general mathematical achievement along the content of the text. The degree of congruence is higher than the pre-test case and the coefficient is .941. Thus the new matched general factor for teacher ratings is highly congruent with the students' general achievement on the post-tests.

Let us summarize our findings for congruent factors. Only one factor from student post-test data is regarded as congruent with a factor from teacher ratings. The factor coefficients of the items for students' general performance at the end of instruction are similar to the coefficients of the items for a new factor for teacher ratings. Items asking for the understanding of basic mathematical concepts and the skill in algebraic manipulation (based on the contents of the text) are closely related to these two congruent factors (one for student performance and one for teacher ratings). In fact the student general achievement factor (the first principal factor for post-test data) is itself maximally congruent with the new factor for teacher ratings. Also, the first principal factor for teacher ratings (response set) is almost maximally congruent with students' general ability.

Another new factor for teacher ratings was matched with student pre-test performance with high congruence, although the degree of the congruence is slightly lower than that obtained for the post-test data. Student ability to work on the tasks which are closely related to the content of the text contribute to this factor, but ability on the tasks which need some knowledge and practice does not contribute. Both of these abilities contribute in the case of post-test data.

Student gains from pre-test to post-test failed to show a high congruence with teacher ratings. We found some positive relationship between mean values of student gains and those of teacher ratings in the earlier analysis. But, in the analysis of factorial structures, in which the effects of mean gains and mean ratings were taken out, a high similarity between both structures was not found.

That is, the strongest relationship between student gain scores and teacher ratings is that between their mean values.

2. Problems and Suggestions for Future Study

In the preceding section, we have analysed and discussed the results we obtained. But still some problems exist in this study. Some of the problems will be discussed and studies needed in the future will be suggested.

First of all, we found that the intercorrelations among student scores of items, particularly, among gain scores were generally lower than the intercorrelations among teacher ratings. Usually, the intercorrelations among cognitive tests are higher than those we have obtained for the student data here.* One reason would be because what we used for correlations were not a set of long tests but a set of short composites of four sub-items. This was done because we intended to prepare and administer as many kinds of items as possible within a given limit of time. If much time is allowed for administration, we should use a longer test which consists of many items, as a unit. This is particularly important for obtaining meaningful gain scores, since the reliability of gain scores will usually be less than the reliabilities of the pre- and post-test scores unless both are highly reliable. (Guilford, 1954, p. 394; Cronbach, 1960, p. 287).

We obtained only two or three common factors from the intercorrelations among student scores in spite of the fact that we used 54 items as variables. They account for only a small portion of the total variance of the variables — about 20% for the pre- and post-tests and 10% for the gain scores. Such low communalities resulted from the low correlations among test scores, and, further, they came from the fact that we used small units of sub-items as variables. This defect was especially evident for the gain scores. The reason why we could not find a fruitful conclusion from the factor analysis of the gain scores would mostly come from this reason.

---

* See the data from French (1951) or Thurstone (1958), for example.

For the factor-analytic model used in this study, we started from the analysis of intercorrelation matrices rather than the analysis of variance-covariance matrices or the cross-product matrices. In the study of the relationship between teacher mean-ratings and student mean-gain scores on items, some positive, not strong, relationship has been found except for a few items. In the factor analysis of the intercorrelations of gain score, this effect is taken out since all the mean values of items are standardized to be zero. With this model, it is only interesting what group of items gained in the same direction and what group of items gained in other directions. If we take the absolute magnitude of gain scores into consideration, some other result might be found.

The similar argument may be possible for the variance-covariance of items. All the variances of items are standardized to unity in the model of this study. The units of measurement for student scores and teacher ratings are different and the variability of the samples of subjects are also different. This problem has not been deeply considered here. For future study, some modified model in which these effects are taken into account would be necessary.

On the construction of test items as stimuli, the items were collected and made in a somewhat arbitrary way. Since this kind of study was the first trial for the UICSM subject material, we did not assume, a priori, any definite dimensions of variables. This study was rather an exploratory type for finding what dimensions are important to student performance and teacher ratings, and no rigorous experimental design was made assuming the important dimensions of factors, as experimental psychologists and statisticians are likely to do. It would be necessary in a future study, however, that the stimulus materials be arranged from this point of view on the basis of the finding from the factor analytic study.

Several variations are possible concerning the intervals between the pre-test and post-test. In this study, Test I based on the first three chapters of the text was pre-tested when teaching for Chapter 1 started and it was post-tested when teaching for Chapter 3 ended. Test II based on Chapters 4 and 5 of the text was pre-tested when the teaching for Chapter 4 started and it was post-tested when the teaching for Chapter 5 ended. We could, of course, give both Tests I and II as a pre-test when the teaching for Chapter 1 starts, and give them as a post-test when the teaching for Chapter 5 or the first semester ends. Or, we could split Tests I and II into small sections for each chapter and give them at the intermediate points in each chapter. We could also design an experiment so that both Tests I and II are given at every section of the chapters or some equivalent place in order to find the effects of transfer and forgetting of the previous chapters. Such kinds of research need more careful and long-term experiments, but they seem important to understanding the process of student learning.

The teacher-held objectives should be reflected on the effect of their teaching to their own students. What we want to know is the facts on which the students who have been taught by a teacher who holds such and such objectives showed such and such progress during the study of course. We used only four teachers and their students from one school for this purpose. To draw a fruitful conclusion in this respect, we need more classes from different schools and their teachers. A longer term and larger scale of research is also required.

At any rate, we are standing just at the introductory stage, and this study would serve as the first step to the long way to the ultimate goal.

CHAPTER VI

SUMMARY AND CONCLUSIONS

This study was designed to investigate some possible relationships between teacher-held objectives and student performance that appeared in responses to samples of test items of the new UICSM high school mathematics.

Two test booklets, Test I and Test II, were made on the basis of the contents of the first five chapters of "High School Mathematics, Course 1" by Max Beberman and H. E. Vaughan (1964). Most of the items consisted of four sub-items with the same content and style. The scores were obtained for each item by the number of correct answers of the sub-items. Twenty-five and twenty-nine items were used for Test I and Test II, respectively.

Test I, as a pre-test, was given to 9th grade Pekin Community High School students in Illinois when classes started, and the same test was given, as a post-test, to the same students when instruction for the first three chapters of the test on which the test contents were based was finished. Test II, as a pre-test, was also given to the same students on the day after Test I was given as a post-test. Test II, as a post-test, was given to the same students when instruction for Chapters 4 and 5 on which the test contents were based was finished. Each test was given using the ordinary 60 minutes-class period. The gain scores for 154 students were obtained for each item.

A factor-analytic method was applied to the three sets of intercorrelations among items, based on the pre-test, post-test and gain scores of students. Three, two, and two factors were obtained from the pre-test, post-test and gain scores, respectively.

A questionnaire was made asking teachers to evaluate the suitability of the test items (the same as those given to the students) for use in an achievement test at the end of instruction in given chapters. The ratings for items were assumed to be an indirect indication of the teacher-held objectives for teaching the subjects. The questionnaire was sent to all the teachers who were using the new text, and 105 responses were collected from 70 different schools in 19 different states.

The intercorrelations among teacher ratings for items were factor analyzed and five factors were extracted. In order to see the factorial congruence between teacher ratings and student performance, a canonical type of analysis was carried out for the sets of the student pre-tests and the teacher ratings, of the student post-tests and the teacher ratings, and of the student gains and the teacher ratings. The factors found from both student scores and teacher ratings were transformed so that the maximum congruence between them was obtained.

The main results found through the analyses in this study were as follows:

1. No particular relationships were found between mean values of the teacher ratings and the student performance on either pre-test items or post-test items, even when the mean values for Pekin teachers only were considered.

2. A weakly positive relationship was found between mean values of the teacher ratings and the student gains. The teachers emphasized sound understanding of basic concepts rather than the simple rote learning of mathematical concepts or simply mechanical computations. Generally, the student progress meets this kind of teacher objectives, but the students showed the highest progress in numerical type problems for which some practice was required.

3. The first principal factor for the student pre-tests indicated a general aptitude for learning mathematics. The ability to translate a verbal statement into a mathematical expression, for example, was important, while a computational skill itself was not important at this stage.

4. The second principal factor for the student pre-tests was related to the order of the items which the student attempted. Time-consuming items and

the items near the end of the tests were likely to have high coefficients on this factor.

5. The third principal factor for the student pre-tests was hard to interpret, and it was left unidentified.

6. The first principal factor for the student post-tests indicated general achievement for mathematics related to the given contents of the text. The ability which was emphasized in the text was closely related to this factor.

7. The second principal factor for the student post-tests seemed to be related to a deductive-inductive factor. But it was not clear.

8. The factors found among student gain scores were ambiguous and interpretation was considered too tenuous to warrant further consideration.

9. For the principal factors for teacher ratings, the first factor was related to individual differences in the general tendency of each teacher's ratings, i.e., teachers' general response set. Items asking for the understandings of the basic mathematical concepts tended to have high coefficients on this factor, and items which were irrelevant to the test tended to have low coefficients.

10. The second and the third factor were related to the irrelevance of the items to the contents in the text. The second factor was related to simple computational abilities which will be taught later. The third factor, however, referred to abilities something apart from the text. The second factor was also related to the conventional vs. new mathematics problems.

11. The fourth principal factor referred to a conceptual vs. algebraic ability. Some skill of algebraic manipulation such as simplifying mathematical expressions and solving equations was positively related to this factor.

12. The fifth principal factor was difficult to interpret and it was left unidentified.

13. Only one factor each from student pre-test and post-test data was congruent with teacher-rating factors. The congruent factor between student performance in the pre-tests and teacher ratings was related to the knowledge that students already had that was closely related to the contents of the text. Ability which needs practice was not important. The congruent factor between student performance in the post-tests and teacher ratings, however, was more related to the achievement of the objectives which were supposed in the text. Both basic concepts and algebraic manipulation emphasized in the text were important for both students and teachers. The contents which were not taught were naturally unimportant for both students and teachers. It should be noticed that the first principal factor for the student performance on the post-tests was itself most congruent with teacher ratings. The two factorial structures matched with students' general achievement and teachers' ratings were highly congruent (coefficient of congruence, .941).

14. A low coefficient of congruence was obtained between the student gains and the teacher ratings, and no consistent patterns were found between them. The two factorial structures, in which the effect of the mean values of student gains and of teacher ratings were taken out, were no longer similar with each other. Although there was some positive relationship between mean values of student gains and of teacher ratings, the second matched-factors in all three cases of student performance were also non-congruent with the teacher ratings.

15. Finally several problems included in this study were discussed and some necessary studies in future were suggested.

BIBLIOGRAPHY

Anderson, T. W. *Introduction to multivariate statistical analysis*. New York: Wiley, 1958.

Barlow, J. A. and Burt, C. The identification of factors from different experiments. *Brit. J. statist. Psychol.*, 1954, *7*, 52-56.

Bartlett, M. S. Internal and external factor analysis. *Brit. J. Pshchol.*, *Stat. Sect.*, 1948, *1*, 73-81.

Beberman, M. and Vaughan, H. E. *High school mathematics*, *course 1*. Boston: D. C. Heath, 1964.

Bloom, B. S., (Ed.) *Taxonomy of Educational Objectives*, *Handbook 1*: *Cognitive Domain*. 1956, New York: David McKay Co.

Borgatta, E. F., Cottrell, L. S. Jr. and Meyer, H. J. On the dimensions of group behavior. *Sociometry*, 1956, *19*, 223-240.

Burt, C. Factor analysis and canonical correlations. *Brit. J. Psychol.*, *Stat. Sect.*, 1948, *1*, 95-106.

Burt, C. The factorial study of temperamental traits. *Brit. J. Psychol.*, *Stat. Sec.*, 1948, *1*, 178-203.

Cattell, R. B. *Personality and motivation structure and measurement*. New York: World Book Co., 1957.

Cronbach, L. J. *Essentials of Psychological Testing*. New York: Harper and Brothers, 1960.

Eckart, C. and Young, G. The approximation of one matrix by another of lower rank. *Psychometrika*, 1936, *1*, 211-218.

Eysenck, J. J. *The structure of human personality*. London: Methuen, 1953.

French, J. W. The description of aptitude and achievement factors in terms of rotated factors. *Psychometric* *Monographs*, *No. 5*, Chicago: Univ. of Chicago Press, 1951.

French, J. W. School of thought in judging excellence of English themes. *1961 Proceedings of Invitational Conference on Testing Problems*, Educational Testing Service, 1962, 19-28.

Gibson, W. A. An expansion upon Tucker's inter-battery factor analysis. *Amer. Psychologist*, 1960a *15*, 487.

Gibson, W. A. Remarks on Tucker's inter-battery method of factor analysis. *Psychometrika*, 1960b, *25*, 19-25.

Gibson, W. A. A retraction on inter-battery factor analysis. Psychometrika, 1961, 26, 451-452.

Guilford, J. P. Psychometric Methods. New York: McGraw-Hill, 1954.

Guilford, J. P. The structure of intellect. Psychol. Bull., 1956, 53, 267-293.

Guilford, J. P. and Zimmerman, W. S. Fourteen dimensions of temperament. Psychol. Monogr., No. 70, 1956.

Gulliksen, H. Intrinsic Validity. Amer. Psychologist, 1950, 5, 511-517.

Guttman, L. A new approach to factor analysis, the radex. Lazarsfeld, P. F. Ed., Mathematical thinking in the social sciences, Glencoe: The Free Press, 1954, 258-348.

Harman, H. H. Modern factor analysis. Chicago: Univ. of Chicago Press, 1960.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Horst, P. Generalized canonical correlations and their application to experimental data. J. clin. Psychol. Monogr. Supplement, No. 14, 1961, 332-347.

Horst, P. Matrix algebra for social scientists. New York: Holt, Rinehart and Winston, 1963.

Hotelling, H. Analysis of complex of statistical variables into principal components, J. educ. Psychol., 1933, 24, 417-441, 498-520.

Hotelling, H. The most predictable criterion. J. educ. Psychol., 1935, 26, 139-142.

Kendall, M. G. Rank Correlation Methods, 2nd ed. London: Charles Griffin, 1955.

Leyden, T. The identification and invariance of factors. Brit. J. statist. Psychol., 1953, 6, 119.

Mager, R. F., Preparing Objectives for Programmed Instruction. 1962, San Francisco: Fearson Press.

Meredith, W. Note on factorial invariance. Psychometrika, 1964, 29, 177-185.

Meredith, W. Rotation to achieve factorial invariance. Psychometrika, 1964, 29, 187-206.

Pinneau, S. R. and Newhouse, A. Measures of invariance and comparability in factor analysis for fixed variables. Psychometrika, 1964, 29, 271-281.

Schuessler, K. F. and Driver, H. A factor analysis of primitive societies. Amer. Soc. Rev., 1956, 21, 493-499.

Thomson, G. H. The maximum correlation of two weighted batteries (Hotelling's "most predictable criterion"). Brit. J. Psychol., Stat. Sect., 1947, 1, 27-33.

Thurstone, L. L. Primary mental abilities. Psychometric Monographs, No. 1, Chicago: Univ. of Chicago Press, 1938.

Thurstone, L. L. Multiple factor analysis. Chicago: Univ. of Chicago Press, 1947.

Thurstone, L. L. and Thurstone, T. G. Factorial studies of intelligence. Psychometric Monographs, No. 2, Chicago: Univ. of Chicago Press, 1941.

Tucker, L. R. A method for synthesis of factor analysis studies. Personnel Research Section Report, No. 984, Washington, D. C., Dept. of the Army, 1951.

Tucker, L. R. An inter-battery method of factor analysis. Psychometrika, 1958, 23, 111-136.

Tucker, L. R. Factor analysis of relevance judgments: an approach to content validity. 1961 Proceedings of Invitational Conference on Testing Problems, Educational Testing Service, 1962, 29-38.

Tucker, L. R. Recovery of factors from simulated data. A paper read at the meetings of The Psychonomic Society and The Psychometric Society. Niagara Falls, Ontario, October 10, 1964.

Walker, Helen M. and Lev, J. Statistical inference. New York: Henry Holt, 1953.

Wrigley, C. and Neuhaus, J. O. The matching of two sets of factors. Amer. Psychologist, 1955, 10, 418-419.

Zachert, V. and Friedman, G. The stability of the factorial pattern of aircrew classification tests in four analyses. Psychometrika, 1953, 18, 219-224.

VITA

Hiroshi Ikeda was born in Kurashiki-city, Japan, on April 19, 1932. He received A. B. in Education from the University of Tokyo in 1956, and A. M. in Educational Psychology from the same institution. He was a teaching and research assistant of Educational Psychology at International Christian University, Tokyo, from 1958 to 1961 at a half-time basis. He has been a graduate student in Educational Psychology at the University of Illinois since 1961, and anticipates his Ph. D. degree in June of 1965. He participated in a summer program for graduate students in educational measurement at The Educational Testing Service, Princeton, 1963. He has been a staff member of the UICSM Mathematics Project since 1964.

His publications include "A Study for Evaluation of the Guttman Scale," Educational Study, No. 7, International Christian University, 1960 (in Japanese), "Vector Interpretation of Classical Test Theory," Japanese Journal of Psychology, 1964, 35, 194-203 (in Japanese), "Intraclass Correlation as an Approximation to the Coefficients of Generalizability," (with L. J. Cronbach and R. A. Avner), Psychological Reports, 1964, 15, 727-736, and "A General Principle of Factor Extraction," Japanese Journal of Psychology, 1965, 35, 288-296 (in Japanese).

He is a member of the Psychometric Society, American Statistical Association, Japanese Psychological Association, and Japanese Association of Educational Psychology. He is also a member of Kappa Delta Pi, an honor society in education.

APPENDIX A

TABLES
OF
INTERCORRELATIONS
AMONG
TEST ITEMS

TABLE I

Intercorrelations of Pre-test Scores among Items*

Items	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	252											
3	172	056										
4	-029	-020	456									
5	024	101	035	098								
6	121	-159	-135	-163	-071							
7	075	-042	-170	-131	065	041						
8	149	-033	-070	-044	024	067	299					
9	-010	002	-007	-100	047	072	157	052				
10	-027	-126	058	025	043	-108	070	-127	013			
11	132	111	-138	-147	-028	188	131	104	-016	-033		
12	023	229	045	-063	-014	020	025	-133	-056	113	079	
13	-003	104	023	-019	043	-092	-133	088	-150	-140	084	-006
14	036	082	-007	026	-025	-048	-030	-034	-034	028	025	111
15	130	011	075	-023	-018	-016	-024	046	-064	044	019	146
16	011	041	-062	021	-063	119	015	060	-071	076	230	182
17	117	-007	-108	056	016	-004	431	276	076	025	114	-087
18	-035	-099	-169	-052	000	129	184	257	096	-072	173	009
19	083	095	078	162	073	-031	-063	-111	-100	100	097	-061
20	-011	104	013	165	246	-020	041	068	096	-061	-019	-082
21	007	083	-001	121	135	041	092	-034	123	153	066	080
22	-007	058	-039	033	025	033	048	121	111	-059	101	-057
23	-100	-028	-052	057	094	111	169	056	090	018	030	-122
24	087	165	068	102	046	-045	-080	088	-155	-126	-006	018
25	-099	-006	-073	009	115	-148	-032	126	-053	-002	096	016

*Decimal points are omitted.
Underlined correlations are significantly different from zero at the 5% level.

TABLE I (Continued)

Intercorrelations of Pre-test Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14	298												
15	133	216											
16	-083	022	179										
17	-134	018	-084	-052									
18	073	-052	024	031	278								
19	131	084	088	-001	009	005							
20	012	129	-096	-106	-017	128	188						
21	-133	053	-226	080	023	-026	125	275					
22	114	075	031	054	-002	141	245	216	204				
23	131	023	-067	003	007	090	218	232	164	535			
24	-031	-151	-112	-099	-025	053	145	247	167	310	290		
25	-004	-078	-051	004	091	150	010	079	096	093	156	276	

TABLE I (Continued)

Intercorrelations of Pre-test Scores among Items

Items	1	2	3	4	5	6	7	8	9	10	11	12
26	054	095	026	080	-030	-049	031	047	068	152	-066	059
27	072	056	110	138	095	-026	-038	064	-024	197	037	056
28	026	050	-007	-039	074	063	-023	-055	011	207	060	032
29	102	033	187	107	096	-159	198	191	031	-018	048	-017
30	178	041	282	214	076	022	-052	-113	-002	119	077	-043
31	-047	-145	106	114	042	012	-036	-047	-072	100	-048	-125
32	-114	082	216	025	-037	-113	002	-074	045	061	041	143
33	-065	079	-008	099	-128	-093	012	021	087	038	139	119
34	-084	-182	003	035	007	051	101	056	-009	-048	037	035
35	227	271	147	050	119	-149	022	151	080	169	139	249
36	054	089	007	-027	-035	008	143	-006	093	230	068	299
37	109	129	150	170	060	-115	143	175	-021	071	031	153
38	065	053	184	031	123	053	028	-021	183	087	017	243
39	270	234	136	046	193	014	054	127	101	159	004	008
40	243	077	216	061	178	087	-074	-019	-013	073	-003	136
41	-105	-047	-064	068	110	068	-069	-096	-118	130	-042	102
42	262	175	105	044	125	070	080	049	151	004	121	156
43	129	168	166	034	018	-005	132	036	010	070	129	156
44	140	200	144	088	004	-068	-013	111	-083	019	021	098
45	-012	-022	098	178	036	-035	-021	-055	071	-025	017	025
46	075	075	214	210	153	-102	061	029	015	-034	-031	068
47	-049	009	098	138	085	-016	016	086	118	-027	-050	-063
48	-169	-020	022	123	039	030	-003	-047	105	062	016	064
49	044	018	068	188	034	049	098	-050	123	-041	092	000
50	-038	016	037	108	-163	041	170	048	009	-052	031	-020
51	043	091	-071	075	020	-069	183	071	073	075	-003	-055
52	-005	116	058	167	139	-0[illegible]8	047	180	014	-023	-177	-092
53	108	032	-136	-044	072	214	080	158	057	007	063	064
54	073	080	-051	025	123	-179	041	090	096	090	046	-043

TABLE I (Continued)

Intercorrelations of Pre-test Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
26	061	019	-028	-037	-077	-153	080	181	093	114	113	109	052
27	118	032	126	046	-080	-072	062	130	167	029	043	254	103
28	-037	-005	-033	065	077	-138	-105	-004	131	005	054	-014	-083
29	076	-120	113	-068	040	050	013	096	008	043	022	047	043
30	-110	075	128	150	-011	-190	203	189	-021	035	038	106	044
31	078	059	003	-020	114	-039	-031	011	024	056	005	016	033
32	-052	063	058	133	-126	-039	073	044	089	-049	-061	040	033
33	-010	-027	-069	219	026	052	-042	-013	148	008	063	028	049
34	109	-055	-063	136	-008	134	-204	-013	-022	-078	-078	-056	114
35	-052	-065	062	127	068	-088	060	-095	-088	-035	-039	080	091
36	004	156	020	092	053	-030	096	072	137	074	215	-018	-003
37	012	018	090	149	026	038	-049	-091	-048	-120	-048	026	043
38	-054	-009	037	034	139	-055	-002	109	076	-035	-011	048	032
39	-014	073	-014	066	-034	-025	113	079	102	090	051	199	021
40	017	153	103	037	044	-076	103	040	011	-042	-042	066	031
41	-085	-071	105	-059	-070	-004	019	036	-032	-022	014	-011	-079
42	-040	113	073	063	067	-023	035	116	106	057	073	048	020
43	-034	-008	060	205	009	-087	088	-012	131	179	049	142	-001
44	049	005	022	037	-064	-104	-047	056	-020	097	083	086	-019
45	-114	-067	036	051	-176	-163	050	210	095	162	081	143	-054
46	060	-018	-078	-025	193	-084	089	105	189	146	145	142	-009
47	100	011	-094	-032	-068	-017	014	265	069	142	042	189	062
48	038	020	-096	040	-040	-036	089	144	028	242	110	170	-027
49	-051	-011	-023	080	050	-182	110	009	066	017	-018	000	-103
50	031	-079	058	109	-060	094	027	-005	-022	054	185	035	166
51	043	-095	006	-016	052	125	091	092	074	220	209	172	022
52	041	-065	-017	-130	-020	024	019	160	038	173	192	300	126
53	026	-115	-027	-122	031	251	085	-019	033	077	167	092	102
54	009	-033	-022	020	-077	006	151	030	152	179	256	159	059

TABLE I (Continued)

Intercorrelations of Pre-test Scores among Items

Items	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
26															
27	180														
28	165	190													
29	-069	179	-166												
30	079	226	242	011											
31	-037	140	102	052	220										
32	037	192	-024	039	232	001									
33	014	168	142	061	243	025	140								
34	029	027	048	044	066	323	-079	132							
35	143	269	099	107	211	-039	083	238	020						
36	329	043	225	-074	049	008	050	137	108	204					
37	104	191	058	204	115	-159	060	356	128	420	176				
38	114	268	175	070	190	-065	118	227	086	232	194	201			
39	255	069	009	-015	142	006	082	-039	-029	165	147	171	068		
40	161	146	078	-018	113	179	082	022	-027	247	111	048	219	348	
41	010	042	-041	-045	-188	-054	-071	-069	-019	-009	-041	000	069	-116	049
42	048	093	141	042	190	-060	080	075	-046	314	181	237	134	136	208
43	120	081	088	113	151	-091	101	141	-026	213	176	189	134	099	280
44	103	104	063	087	108	-170	-099	004	-130	174	125	050	-014	060	-059
45	047	089	076	100	209	028	061	128	-024	099	094	041	031	068	002
46	070	163	066	176	039	137	125	087	032	221	242	074	179	019	048
47	065	212	038	094	191	127	185	-043	006	044	184	-015	110	166	044
48	024	129	073	-037	192	101	061	-014	007	043	184	-070	063	128	034
49	135	203	089	086	170	-029	154	091	-041	095	037	159	029	041	000
50	055	-042	-009	126	100	-024	066	203	158	038	216	199	031	069	-051
51	153	139	-093	119	066	090	051	005	048	100	099	028	-200	264	-031
52	190	188	-130	111	-063	-009	042	-068	-054	001	-062	047	-041	313	144
53	009	077	-128	241	-141	-084	-120	-059	043	068	075	018	044	-020	-061
54	073	282	-006	230	-008	-019	-043	032	-122	087	109	018	020	062	-029

TABLE I (Continued)

Intercorrelations of Pre-test Scores among Items

Items	41	42	43	44	45	46	47	48	49	50	51	52	53	54
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42	-054													
43	-028	407												
44	-201	148	126											
45	-110	185	211	220										
46	-025	122	164	074	143									
47	-105	215	094	196	260	296								
48	-159	128	107	191	325	288	701							
49	-205	165	181	144	106	114	038	039						
50	-144	-033	058	123	074	156	084	058	112					
51	-026	135	118	068	247	089	129	241	039	217				
52	-051	-049	103	154	122	021	218	227	188	-014	351			
53	-040	-072	-015	016	-116	050	-054	045	040	173	277	196		
54	-068	076	100	164	097	226	195	236	148	120	352	237	330	

TABLE II

Intercorrelations of Post-test Scores among Items*

Items	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	267											
3	206	170										
4	166	005	123									
5	200	097	-002	-010								
6	-025	023	004	-015	084							
7	099	-097	097	084	021	-114						
8	037	-122	006	026	067	-119	275					
9	-041	-033	-039	115	054	-070	056	-014				
10	-084	022	117	-014	-006	060	-013	-087	171			
11	206	327	074	078	024	099	010	013	-094	073		
12	188	142	013	009	129	-075	197	169	-025	067	216	
13	128	042	-060	-023	124	-004	001	-061	015	-041	106	258
14	171	093	-006	-020	189	-073	086	134	096	-104	129	192
15	-059	044	-027	-042	103	-167	101	106	078	071	130	125
16	033	023	016	-166	068	011	110	083	-030	072	-069	108
17	-009	-098	085	-083	030	-088	171	090	-011	081	084	078
18	143	075	047	-080	194	-009	147	-015	-027	117	043	131
19	107	168	170	-021	114	-222	-050	037	098	094	019	026
20	181	-039	033	060	162	-074	-008	076	152	082	-028	032
21	158	095	027	115	-019	136	082	-031	042	-029	-123	108
22	167	-063	080	038	119	-156	098	047	157	-032	019	107
23	086	154	-025	036	-045	-173	108	061	166	083	-030	056
24	190	090	067	113	062	-166	150	108	070	-023	034	266
25	092	063	-035	-015	091	035	-026	005	223	032	055	016

* Decimal points are omitted.
Underlined correlations are significantly different from zero at the 5% level.

TABLE II (Continued)

Intercorrelations of Post-test Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14	476												
15	-012	029											
16	048	041	065										
17	-025	024	027	052									
18	146	066	010	078	552								
19	079	237	080	-060	166	142							
20	182	131	078	168	141	185	106						
21	-005	-014	-070	-092	036	-041	098	121					
22	071	141	047	-147	006	021	067	113	121				
23	-032	060	089	058	003	015	108	007	044	014			
24	101	182	172	031	078	160	209	149	080	113	146		
25	066	060	026	116	-067	097	-002	167	-153	099	063	147	

TABLE II (Continued)

Intercorrelations of Post-test Scores among Items

Items	1	2	3	4	5	6	7	8	9	10	11	12
26	043	058	005	061	025	052	084	002	-037	074	025	124
27	103	206	058	007	217	-030	025	203	052	053	059	153
28	286	199	177	086	076	-024	018	086	016	056	001	165
29	174	038	023	126	144	-182	166	171	212	059	007	164
30	280	-006	-068	-057	085	002	039	124	088	-127	096	247
31	152	124	136	094	082	127	030	078	136	056	037	273
32	079	-061	058	065	008	010	081	182	044	001	-094	107
33	197	155	090	-018	193	043	004	135	-039	-033	-025	234
34	156	067	105	-051	176	-223	-027	032	097	157	-026	168
35	106	143	132	-050	003	-021	119	283	-072	-073	031	350
36	076	027	-140	-012	-027	-095	152	205	032	-022	061	156
37	028	169	158	-080	-008	-013	057	156	082	010	-025	308
38	189	071	122	015	-103	-057	103	148	-012	-018	-045	255
39	155	-007	-017	128	134	-012	077	222	-007	-105	033	174
40	163	150	105	136	160	-106	102	137	075	001	196	258
41	025	-062	-156	-108	-060	137	-024	-030	018	049	-037	-098
42	212	101	-019	127	023	087	065	236	012	-019	210	109
43	179	205	051	-003	115	058	-033	145	112	188	164	219
44	219	-038	126	067	-016	029	116	146	162	-039	-012	123
45	033	054	125	016	142	008	243	216	055	084	-058	245
46	189	197	025	026	076	081	056	052	-053	043	070	126
47	253	181	086	047	037	059	074	028	012	138	182	229
48	025	146	032	014	011	069	049	087	-020	-016	141	019
49	206	144	041	068	026	-042	068	142	003	009	084	138
50	025	193	035	073	126	-073	100	097	146	159	-016	257
51	-064	023	-054	-022	-003	051	235	237	029	-051	063	019
52	142	032	-045	013	071	-061	096	155	072	031	097	172
53	-135	151	-024	-009	-032	033	027	106	-050	-052	-018	-081
54	-081	152	-102	-013	-016	090	113	130	135	-021	-098	-086

TABLE II (Continued)

Intercorrelations of Post-test Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
26	205	222	-050	013	-061	212	163	053	085	027	132	052	081
27	-017	266	116	-032	092	108	168	089	095	181	103	135	030
28	-008	147	059	-008	-148	062	114	158	187	123	188	252	131
29	177	211	084	111	057	152	138	245	-063	145	272	143	191
30	091	282	059	079	039	061	098	271	062	151	040	199	156
31	128	046	-012	026	-002	185	021	-004	-012	082	125	060	069
32	084	194	-163	024	088	112	141	173	050	082	-026	144	084
33	144	169	084	-007	029	129	200	211	119	109	178	060	029
34	114	110	042	133	031	136	137	251	020	160	109	041	214
35	096	244	148	134	092	147	162	188	099	-017	125	181	210
36	155	124	104	049	041	226	065	114	065	092	147	113	110
37	174	238	042	174	023	107	078	049	-041	054	112	183	105
38	148	064	-015	029	-017	078	127	095	226	158	152	134	105
39	169	193	-074	-002	049	166	-020	062	-014	154	-021	108	110
40	215	226	110	003	152	158	122	237	089	173	035	189	081
41	209	088	-018	015	-044	025	-087	097	110	-079	-045	-093	073
42	085	210	-034	007	159	166	179	259	180	091	040	201	107
43	240	345	047	139	060	184	172	147	084	120	089	229	191
44	106	147	037	058	025	127	033	225	134	322	-015	213	171
45	162	258	-052	134	065	114	107	125	171	133	029	168	104
46	091	138	-087	009	-055	113	084	189	026	104	047	188	145
47	091	121	054	060	116	159	062	121	147	174	059	201	034
48	-004	-015	-032	028	033	069	076	089	014	281	-040	057	020
49	089	100	008	-135	-067	050	121	-003	032	131	-012	-059	061
50	137	265	121	038	000	167	234	196	007	168	163	206	088
51	038	114	083	-051	154	160	-054	069	021	102	-044	159	039
52	171	263	025	-058	059	129	110	095	025	114	039	102	052
53	-029	141	109	-066	-131	-082	-055	012	084	069	041	003	017
54	-077	056	071	-056	064	113	002	-011	069	037	083	-003	016

TABLE II (Continued)

Intercorrelations of Post-test Scores among Items

Items	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
26															
27	214														
28	173	335													
29	105	222	278												
30	080	269	209	260											
31	073	201	274	300	282										
32	022	152	222	188	334	379									
33	166	233	273	316	277	350	297								
34	091	225	218	388	203	237	313	315							
35	291	304	403	237	265	274	175	461	185						
36	209	056	114	268	155	159	127	173	019	295					
37	185	203	190	235	274	374	242	322	156	424	261				
38	147	162	261	181	258	192	340	355	273	308	118	134			
39	100	237	253	171	268	246	107	107	101	306	184	160	074		
40	125	203	159	206	169	141	248	237	166	288	163	210	218	165	
41	124	-143	-070	-032	107	083	184	036	015	101	288	068	110	-011	098
42	144	245	167	317	181	101	206	263	180	349	331	153	153	138	328
43	245	279	207	260	217	204	171	256	220	364	248	304	177	210	358
44	091	197	201	234	219	255	287	145	198	212	257	248	238	159	306
45	230	298	148	184	105	171	219	244	182	374	166	317	284	157	258
46	270	244	212	150	192	207	307	205	175	280	282	231	329	084	043
47	174	221	153	235	256	291	177	297	151	292	299	318	164	120	352
48	091	062	100	042	155	189	177	127	146	122	101	036	202	032	085
49	117	075	064	107	176	189	034	115	-001	201	163	170	106	148	100
50	248	283	251	242	299	247	191	243	248	301	287	378	201	152	290
51	024	197	-008	106	154	095	201	067	142	046	049	-013	107	073	202
52	101	183	042	253	261	222	344	288	364	142	116	054	193	220	283
53	-022	153	-028	049	130	002	096	088	193	-004	023	044	090	-000	081
54	045	024	-093	070	065	089	111	085	012	079	241	189	014	072	-019

TABLE II (Continued)

Intercorrelations of post-test Scores among Items

Items	41	42	43	44	45	46	47	48	49	50	51	52	53	54
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42	046													
43	255	394												
44	094	454	307											
45	246	258	360	263										
46	188	326	322	317	400									
47	121	458	299	396	306	294								
48	041	246	239	183	225	346	385							
49	069	267	110	260	153	181	257	248						
50	144	238	367	189	292	313	398	233	176					
51	062	246	138	179	244	163	286	428	179	311				
52	127	346	269	235	232	147	253	179	297	328	464			
53	136	120	175	036	138	095	022	250	135	219	388	387		
54	127	096	106	051	120	193	139	200	173	219	302	210	420	

TABLE III

Intercorrelations of Gain Scores among Items*

Items	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	130											
3	-002	053										
4	-068	-020	358									
5	035	133	-166	-055								
6	135	-148	-012	-057	029							
7	046	-144	-043	-105	010	-069						
8	-009	-062	-081	-046	071	-096	384					
9	014	-200	-087	-051	025	-025	140	005				
10	045	-069	114	-024	-120	089	-034	-044	-044			
11	038	103	-140	-079	-134	142	027	058	-060	-046		
12	-173	055	010	-068	-059	-027	067	-103	-007	080	-044	
13	021	023	-072	-080	-020	-039	-131	-039	038	-150	-099	-076
14	-098	008	-069	-042	022	-012	044	085	134	-086	007	-024
15	-087	008	031	-075	-002	-138	043	091	-004	044	-017	149
16	-050	001	-033	-017	-062	075	-061	-012	-025	119	146	141
17	-039	-080	-097	-049	081	-056	264	150	-017	055	009	091
18	055	-102	-187	-117	174	008	169	114	007	-018	021	149
19	077	160	-008	086	045	-019	-090	-078	-051	149	106	-014
20	-005	-034	-085	-007	142	046	-066	065	148	-089	-037	-054
21	-017	084	-098	092	-032	092	012	-052	131	012	-027	042
22	182	-090	-064	021	083	084	-017	117	117	-072	119	026
23	077	-011	-045	034	-037	009	081	092	156	-039	-004	-006
24	068	007	024	062	-090	-137	-011	057	-031	-096	-027	085
25	-023	-068	-097	-042	000	-132	-025	053	100	-037	031	-050

* Decimal points are omitted.
Underlined correlations are significantly different from zero at the 5% level.

TABLE III (Continued)

Intercorrelations of Gain Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14	439												
15	020	037											
16	-111	-012	139										
17	-034	055	015	007									
18	170	094	076	010	399								
19	112	149	071	-014	094	114							
20	026	035	-020	-006	-036	049	-042						
21	-085	-066	-090	080	-069	-150	067	175					
22	130	073	-070	014	-034	083	158	130	174				
23	129	053	-080	-057	-012	026	183	004	010	389			
24	-016	-068	-032	-052	-057	023	070	117	031	240	209		
25	043	016	-186	006	-070	098	-034	159	047	141	179	259	

TABLE III (Continued)

Intercorrelations of Gain Scores among Items

Items	1	2	3	4	5	6	7	8	9	10	11	12
26	069	059	096	110	-025	-056	054	-001	024	130	-169	004
27	-113	-011	-054	-045	089	-016	-047	122	-116	180	-060	-057
28	048	016	003	-038	-119	057	044	-095	053	187	036	-134
29	068	-197	-018	029	077	-219	243	193	023	-084	-016	011
30	021	-182	016	-066	-045	174	079	-069	040	-056	-004	-059
31	052	-128	-153	-083	024	118	083	062	-054	-004	-043	050
32	-096	-046	-025	-106	-134	-029	122	103	-040	-102	002	050
33	-087	-056	-175	-015	-017	-002	-056	-025	-096	-138	-026	025
34	046	-063	-152	-105	-002	-107	-053	101	-043	004	-086	080
35	-157	124	051	-064	086	-012	043	064	-052	026	065	068
36	-147	-060	-096	-015	-087	-085	184	062	012	131	056	241
37	-160	-011	075	095	-036	047	214	043	-083	002	-035	011
38	066	079	108	028	-093	-000	016	055	-003	-128	-087	-008
39	070	028	022	061	136	-017	021	149	168	156	-003	-115
40	-120	-018	-039	-103	212	012	-062	-010	057	081	-083	087
41	-053	-025	-022	-001	061	044	-082	-023	007	042	-059	-100
42	071	103	-057	-006	139	018	060	069	-095	-034	179	063
43	006	-016	026	002	144	022	008	102	035	050	020	046
44	067	-007	054	024	-044	-040	115	098	000	-005	033	117
45	-063	-049	163	022	061	068	083	-056	029	-024	-009	109
46	089	045	126	095	022	-010	076	-025	-198	-100	-011	051
47	012	-041	-017	055	028	021	025	038	-011	069	-002	033
48	002	-079	-040	-016	062	084	-058	043	-006	-003	013	042
49	097	-019	041	141	077	068	104	051	029	-026	099	-042
50	-038	-100	027	-026	-145	005	181	069	060	003	-005	018
51	-022	018	-044	-103	013	-165	104	104	038	-090	-090	-051
52	026	-088	-066	-067	137	031	071	245	017	-021	-185	076
53	-040	-039	-127	-033	080	116	-093	174	036	-108	013	-131
54	-109	020	-020	110	046	-020	-084	033	-013	-003	117	-005

TABLE III (Continued)

Intercorrelations of Gain Scores among Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
26	062	026	-073	030	-154	-054	007	034	014	-035	081	055	016
27	-140	041	160	006	032	-081	051	049	057	010	075	107	068
28	-134	-055	-009	-103	-091	-120	-040	-018	150	059	086	060	-035
29	025	064	-041	-098	023	096	061	078	-035	055	144	003	101
30	-108	015	-084	-031	-008	-153	-031	154	142	056	-060	059	021
31	119	060	-127	-095	065	155	108	-031	-021	148	178	050	075
32	-014	101	-137	020	088	187	-003	044	034	-003	-079	044	-031
33	002	068	-055	-019	-105	021	-004	-021	008	031	120	-126	-059
34	039	098	-014	084	040	207	-035	124	037	091	034	-020	256
35	-200	-197	084	108	-077	-037	-072	-092	011	-207	-107	-010	057
36	-072	075	042	073	063	049	-048	-103	-036	-054	175	-069	-081
37	-062	086	068	149	-030	086	-094	-139	-168	-184	-010	-070	-063
38	-017	-089	032	-073	033	-059	-131	016	086	-071	-075	-002	-099
39	-079	-031	-040	025	092	175	090	-033	-023	-044	017	152	075
40	-099	-013	113	-004	120	062	082	134	046	-109	-077	153	025
41	001	014	016	-091	-015	030	-026	186	023	-032	-028	-065	-095
42	-085	-023	161	019	053	076	045	073	038	014	-132	022	-026
43	044	160	019	205	059	163	039	130	101	147	-002	160	083
44	-010	001	015	051	-006	099	-083	013	051	103	106	132	046
45	063	031	069	036	008	042	-019	083	137	065	-089	108	011
46	016	-107	-154	-011	-005	-089	-102	-024	-011	040	168	016	-116
47	-069	-082	-095	-024	076	065	034	185	075	074	011	236	-074
48	-094	-096	-239	-007	219	-032	-011	173	010	131	-047	128	-064
49	-050	-033	-017	-075	-090	-006	061	-035	009	-003	-027	-059	-128
50	044	159	082	019	031	139	089	006	-122	-022	099	044	044
51	022	-005	-022	-121	181	144	054	014	-043	054	040	140	019
52	063	133	020	-114	055	167	-028	051	-062	048	031	088	042
53	-042	068	003	-153	-090	026	011	042	074	-114	-052	-129	043
54	-092	051	068	-006	012	-094	039	-047	086	-076	037	-124	016

TABLE III (Continued)

Intercorrelations of Gain Scores among Items

Items	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
26															
27	001														
28	147	162													
29	-073	001	-139												
30	-020	172	141	-040											
31	-008	024	122	202	177										
32	-126	-015	052	019	192	214									
33	-002	-063	-002	099	084	128	156								
34	-031	-068	-033	238	-001	218	180	112							
35	108	066	-095	-021	014	-031	-086	151	-046						
36	342	-006	128	007	-020	053	-029	061	-061	001					
37	117	022	-035	079	-086	013	-021	172	-061	132	140				
38	-159	085	-002	-071	-005	-047	142	085	-042	036	-117	-073			
39	176	040	003	-085	-043	030	011	-191	-061	081	104	-005	-169		
40	-006	064	028	-041	-087	046	063	-154	028	061	-077	-059	124	082	
41	003	-037	-097	-044	-029	-033	068	052	-094	063	013	-010	082	-146	120
42	038	-016	-049	-006	-014	-098	-059	002	-006	126	157	045	-021	070	075
43	120	074	-075	102	029	-019	031	-054	068	-074	036	097	-066	009	231
44	-100	-028	-001	005	-004	081	000	-001	152	-032	-018	045	052	-102	-025
45	-068	066	-105	033	048	025	084	027	-037	063	-138	071	066	-055	032
46	078	017	-050	114	-039	051	104	053	-020	006	125	048	209	-146	-034
47	030	135	069	-108	229	143	152	-055	-041	-059	151	-013	036	068	143
48	-049	034	015	-193	190	082	104	-092	-020	-100	017	-153	-033	093	055
49	103	-074	066	-011	111	092	-118	-038	-103	025	003	065	-051	-013	-089
50	028	071	-071	142	162	187	094	041	230	046	045	152	005	-015	140
51	-001	031	-151	068	072	-035	093	-156	077	007	040	-063	-212	246	021
52	-001	032	-117	154	101	190	209	068	286	-013	-104	-088	-046	186	098
53	-160	059	-128	175	-025	016	018	-021	194	134	-106	030	042	-073	081
54	-021	110	-215	040	098	-036	072	017	-010	072	012	028	062	048	127

TABLE III (Continued)

Intercorrelations of Gain Scores among Items

Items	41	42	43	44	45	46	47	48	49	50	51	52	53	54
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42	071													
43	076	277												
44	-029	178	146											
45	157	010	266	133										
46	050	005	027	094	177									
47	081	164	082	084	203	220								
48	-113	013	-058	093	069	185	555							
49	-060	-064	-021	047	024	104	100	134						
50	001	-066	145	106	003	056	171	-064	123					
51	-134	-086	-014	015	077	-090	021	159	031	192				
52	-002	-129	049	093	092	-060	046	060	123	203	306			
53	-022	-052	-053	034	-075	-129	-177	-118	087	147	127	367		
54	088	-026	003	-010	-058	105	-120	-093	051	205	162	181	345	

TABLE IV

Intercorrelations of Teacher Ratings of Items*

Items	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	304											
3	110	357										
4	100	333	689									
5	210	402	373	413								
6	078	151	193	206	421							
7	093	454	532	562	523	190						
8	080	373	606	617	474	260	750					
9	-016	235	342	347	213	033	513	399				
10	-015	193	278	354	276	028	441	307	750			
11	295	499	402	394	454	172	624	543	525	390		
12	223	391	189	345	233	-002	289	319	331	386	536	
13	-068	290	238	314	132	-215	418	386	478	441	481	603
14	-067	198	278	377	109	007	273	399	418	473	367	593
15	216	275	272	338	378	212	342	408	231	264	312	305
16	320	303	161	288	301	172	342	327	414	424	449	541
17	-093	289	169	340	081	-125	429	334	540	589	360	428
18	-082	196	197	342	032	-113	346	302	519	598	286	494
19	294	345	190	223	310	263	303	291	359	297	389	424
20	316	284	172	264	224	242	214	174	329	249	322	364
21	301	184	062	241	129	163	114	107	220	164	152	342
22	306	314	243	195	106	227	190	225	304	227	249	352
23	265	373	253	233	344	269	297	234	349	260	405	323
24	253	358	516	493	627	364	564	552	306	319	524	377
25	350	119	125	146	186	410	128	113	039	081	160	130

* Decimal points are omitted.
Underlined correlations are non-significantly different from zero at the 5% level.

TABLE IV (Continued)

Intercorrelations of Teacher Ratings of Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14	745												
15	317	301											
16	373	428	543										
17	623	539	267	331									
18	622	585	315	344	868								
19	248	334	405	531	345	374							
20	244	353	435	587	301	334	765						
21	132	176	287	409	233	230	597	682					
22	229	259	229	414	228	345	610	577	567				
23	188	233	365	398	296	312	600	619	469	548			
24	193	256	454	442	103	108	331	293	136	237	433		
25	-169	-017	168	212	002	-077	296	313	417	146	381	301	

TABLE IV (Continued)

Intercorrelations of Teacher Ratings of Items

Items	1	2	3	4	5	6	7	8	9	10	11	12
26	327	295	187	260	548	128	397	345	172	163	412	290
27	356	226	226	357	468	292	366	444	186	128	405	238
28	346	177	208	330	391	282	260	268	069	-024	329	189
29	403	309	368	479	509	249	438	466	131	192	494	290
30	324	347	398	432	520	249	449	452	312	314	503	341
31	237	425	443	540	601	232	569	638	216	249	555	313
32	194	291	270	307	387	090	374	397	309	325	408	343
33	234	348	376	540	509	188	425	521	222	280	437	298
34	211	380	436	518	392	217	475	502	359	340	440	367
35	451	393	281	251	503	206	391	391	292	147	581	295
36	362	393	291	320	527	184	409	461	306	129	614	322
37	437	480	314	310	501	124	483	381	346	205	593	325
38	201	410	331	371	419	316	515	420	401	218	578	389
39	330	483	250	273	509	204	531	395	329	240	504	300
40	376	554	308	385	615	262	527	407	359	287	568	419
41	251	492	403	287	628	070	562	490	339	272	582	285
42	222	496	400	305	676	141	599	521	407	316	547	328
43	227	435	318	385	693	348	571	537	294	321	497	375
44	329	387	209	223	559	339	511	439	209	247	473	319
45	379	434	214	272	547	251	430	390	162	157	440	318
46	455	322	161	199	452	265	276	275	205	153	404	296
47	163	368	295	261	386	237	408	302	435	389	450	293
48	276	275	216	211	395	378	320	181	230	162	394	174
49	303	248	242	322	376	368	161	242	130	064	221	235
50	358	218	453	375	380	342	222	301	133	097	326	130
51	299	595	482	544	658	278	695	598	319	316	643	429
52	287	519	453	505	600	346	621	543	331	300	568	352
53	102	250	275	273	204	240	253	393	366	204	317	369
54	218	394	283	306	411	363	489	414	335	307	482	299

TABLE IV (Continued)

Intercorrelations of Teacher Ratings of Items

Items	13	14	15	16	17	18	19	20	21	22	23	24	25
26	204	159	418	321	182	125	368	268	220	077	327	441	270
27	125	231	451	379	209	101	326	356	244	191	355	418	290
28	-026	025	269	323	-048	-069	324	401	410	179	342	402	412
29	223	255	491	423	182	193	360	365	219	197	337	590	284
30	198	263	467	483	282	223	414	405	322	247	488	543	349
31	288	282	436	357	242	168	248	179	038	053	312	600	141
32	324	323	353	505	280	202	314	303	148	106	370	456	181
33	309	331	416	371	294	218	243	212	124	155	327	539	130
34	448	497	366	454	380	374	447	384	267	355	283	389	078
35	196	158	227	335	122	014	355	282	205	135	316	425	316
36	294	218	251	294	179	072	336	263	169	110	310	456	304
37	178	128	239	342	228	115	358	290	272	192	349	414	285
38	233	221	287	435	270	162	539	434	416	293	443	470	292
39	273	183	432	310	222	101	364	432	395	193	366	450	274
40	305	266	407	459	330	222	484	417	300	318	457	576	223
41	298	164	354	295	261	144	280	167	042	087	301	457	108
42	300	195	425	386	251	165	333	226	088	156	360	573	104
43	235	190	390	364	203	144	308	269	210	196	327	573	179
44	155	137	319	292	115	044	348	244	189	176	268	453	373
45	210	146	381	432	195	132	422	308	222	261	361	403	280
46	146	105	308	391	149	085	437	351	256	194	308	425	386
47	241	187	097	272	222	167	454	288	210	241	311	456	278
48	038	020	147	181	052	029	426	348	256	259	375	401	420
49	-017	168	348	279	058	028	266	380	337	247	265	417	428
50	035	180	291	202	045	022	234	360	307	183	355	310	421
51	353	279	363	445	288	188	347	292	162	211	341	650	219
52	274	262	353	434	283	241	371	294	178	288	328	589	209
53	220	316	232	318	329	328	491	406	517	453	386	200	202
54	112	260	198	234	152	130	398	262	262	302	296	471	313

TABLE IV (Continued)

Intercorrelations of Teacher Ratings of Items

Items	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
26															
27	563														
28	423	437													
29	572	639	578												
30	610	623	491	712											
31	510	590	399	713	616										
32	431	487	367	510	590	645									
33	468	598	345	681	586	799	567								
34	332	441	288	554	482	586	436	673							
35	579	552	517	476	503	515	518	401	343						
36	532	506	468	483	480	576	513	498	421	888					
37	550	511	423	431	479	553	486	470	362	789	755				
38	460	385	526	409	492	453	371	388	480	503	530	536			
39	444	452	419	448	480	439	386	412	340	531	526	594	520		
40	524	573	411	527	554	517	526	522	454	622	626	636	530	713	
41	514	473	303	502	551	554	403	514	393	648	651	646	381	542	604
42	476	531	291	495	556	574	466	530	425	613	631	616	498	596	686
43	470	466	400	464	490	570	393	514	422	522	569	521	548	652	637
44	505	472	396	413	403	486	313	380	316	616	593	559	524	569	562
45	444	495	288	477	490	468	422	418	359	578	562	568	398	444	542
46	375	413	331	375	429	368	315	367	277	511	513	518	317	427	506
47	353	276	276	307	440	300	407	337	325	522	504	487	500	467	560
48	347	291	362	360	421	240	211	286	215	473	450	430	442	454	482
49	244	399	339	350	443	288	217	370	194	267	310	356	262	376	328
50	267	442	482	432	536	315	353	377	369	528	495	413	384	467	352
51	490	523	339	593	581	692	466	675	552	613	626	606	494	611	689
52	419	514	317	546	561	585	430	592	616	534	556	532	532	543	659
53	212	273	278	147	398	189	230	182	400	323	308	309	571	202	314
54	325	325	305	410	382	453	295	425	434	398	433	522	543	394	457

TABLE IV (Continued)

Intercorrelations of Teacher Ratings of Items

Items	41	42	43	44	45	46	47	48	49	50	51	52	53	54
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42	889													
43	620	674												
44	627	653	769											
45	612	639	614	612										
46	443	459	490	552	653									
47	437	494	464	500	416	531								
48	334	338	409	475	377	486	791							
49	243	259	399	375	378	568	415	522						
50	381	363	428	361	432	330	371	406	497					
51	686	698	700	592	638	513	516	463	389	471				
52	575	659	695	584	652	436	516	463	365	530	867			
53	213	282	299	302	181	262	365	235	283	349	261	356		
54	397	449	401	477	283	285	501	388	380	327	473	520	461	

APPENDIX B

THE
QUESTIONNAIRE
FOR
TEACHERS

1. Name:________________________________ Male_____ Female_____

2. Circle the numeral corresponding to the chapters in the new UICSM Course 1 test that your most advanced class is now studying.

   1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

3. How many years have you taught mathematics? _____ years since _____.

4. How many years have you taught UICSM Units I and II (previously)? _______ years.
   When did you last teach these units? Fall_____ Spring_____ of _____.

5. Have you taught other UICSM units? Yes_____ No_____
   If Yes, what units have you taught?____________________________

6. Have you taught any subject other than mathematics? Yes_____ No_____
   If Yes, what subjects have you taught? ____________________________

7. How many courses have you taken in which you studied the contents of the UICSM curriculum? ____________________________

8. Did you attend any summer institutes on the UICSM curriculum?
   Yes_____ No_____. If Yes, in what year(s) did you attend?____________

9. Rank the following phrases according to how closely they seem to describe the abilities you stress in teaching UICSM, Course 1. Mark '1' for the most important ability, '2' for the next and so on to the least important ability. Please do not omit any items and do not give a tied-rank.

   _____ Skill in numerical computation

   _____ Skill in symbolic manipulation

   _____ Remembering mathematical principles

   _____ Understanding mathematical concepts

   _____ Discovering mathematical relationships

   _____ Deductive reasoning

   _____ Generalizing from concrete objects to abstract ideas

   _____ Applying mathematical skills to real life problems

10. To help us get an overall picture of how teachers would like to have tests constructed, please mark each of the following scales. Mark each scale to indicate the proportion of items of the two kinds indicated that you would like to see included in Course 1 tests for grading purposes.

For example, mark the first scale to indicate what percentage of relatively easy items, as compared with relatively difficult items, you would like to have included. Mark the second scale to indicate what percentage of multiple-choice items, as compared with completion items, you would like to have. Mark each scale with an 'x' as shown in the sample below.

(Sample)	arithmetic	0% X 50% 100%	algebraic
(a)	relatively easy	0% 50% 100%	relatively difficult
(b)	multiple-choice	0% 50% 100%	completion
(c)	verbal	0% 50% 100%	non-verbal
(d)	many words	0% 50% 100%	few words
(e)	time-consuming	0% 50% 100%	quick response
(f)	straight-forward	0% 50% 100%	tricky
(g)	familiar	0% 50% 100%	unfamiliar
(h)	basic	0% 50% 100%	applied
(i)	concrete	0% 50% 100%	abstract
(j)	computational	0% 50% 100%	conceptual
(k)	reasoning	0% 50% 100%	recall
(ℓ)	inductive	0% 50% 100%	deductive
(m)	teacher constructed	0% 50% 100%	UICSM project constructed

11. Using the scales below, rate all of the items of the enclosed Test I. Indicate your rating for each item by placing an 'x' in one of the 10 boxes of the scale corresponding to the item. Mark the box on each scale which indicates how good the item would be, in your opinion, for inclusion in a test to be given at the end of the first three chapters of the new UICSM text for course one.

Every item except No. 5 and No. 6 has four sub-items. You should ignore differences between the sub-items of a given item and rate each item as a whole.

Please, do not omit any items. If you can describe the reasons for your rating breifly, do so in the space provided at the right of the rating scale.

Item No.	0	1	2	3	4	5	6	7	8	9	Reasons
1.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
2.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
3.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
4.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
5.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
6.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
7.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
8.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________

Item No.	0	1	2	3	4	5	6	7	8	9	Reasons
9.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
10.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
11.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
12.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
13.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
14.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
15.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
16.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
17.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
18.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
19.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
20.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________

Item No.	0	1	2	3	4	5	6	7	8	9	Reasons
21.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
22.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
23.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
24.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
25.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________

. . . .

Overall rating (How good is Test I?)

0	1	2	3	4	5	6	7	8	9
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
worthless		inferior		good		superior			perfect

List the reasons for your overall rating.

12. Using the scales below, rate all of the items of the enclosed Test II. Indicate your rating for each item by placing an 'x' in one of the 10 boxes of the scale corresponding to the item. Mark the box on each scale which indicates how good the item would be, in your opinion, for inclusion in a test to be given at the end of Chapter 4 and 5 of the new UICSM text for course one.

Every item except No. 29 has four sub-items. You should ignore differences between the sub-items of a given item and rate each item as a whole.

Please, do not omit any items. If you can describe the reasons for your rating briefly, do so in the space provided at the right of the rating scale.

Item No.	0	1	2	3	4	5	6	7	8	9	Reasons
1.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
2.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
3.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
4.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
5.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
6.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
7.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
8.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________

Item No.	0	1	2	3	4	5	6	7	8	9	Reasons
9.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
10.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
11.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
12.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
13.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
14.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
15.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
	worthless		inferior		good		superior			perfect	
16.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
17.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
18.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
19.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________
20.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	________

Item No. Reasons

21. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

worthless inferior good superior perfect

22. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

23. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

24. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

25. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

26. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

worthless inferior good superior perfect

27. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

28. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

29. 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ ______

. . . .

Overall rating (How good is Test II?)

0 [ ] 1 [ ] 2 [ ] 3 [ ] 4 [ ] 5 [ ] 6 [ ] 7 [ ] 8 [ ] 9 [ ]

worthless inferior good superior perfect

(continued)

List the reasons for your overall rating.

13. If you have any other comments or suggestions about these tests, please write them below.

APPENDIX C

TEST I

TEST 1

Sample Items
Proposed
for
Chapters 1 - 3
of
High School Mathematics
Course 1

D. C. Heath, 1964

1. **Single quotes, ' ', are used to set off a written or printed symbol (word, numeral, or other expression) about the symbol itself instead of the thing it represents.**

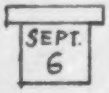
(Examples) Ada is taller than Penelope but 'Ada' is shorter than 'Penelope'.

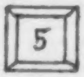
'3' is a symbol that represents or names the number 3.

It makes no sense to say that .0001 is larger than 3.
But '.0001' does take up more space than '3'.

Circle the letter indicating the sentence which makes the most sense.

A B C D (a) A. Mary is a part of Maryland.
B. 'John' has a letter. John has four letters.
C. I have trouble with my pen when I make a 3.
D. He erased the '5' and put a '4' in its place.

A B C D (b) A. 

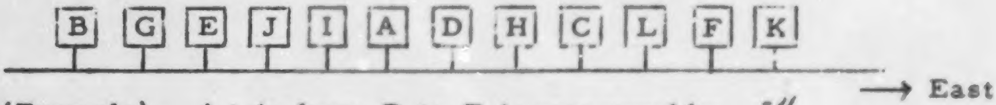
The calender has the number 6 on it.
B. '11 + 3' is NOT the same as '2 × 7'.
C. The number $\frac{32}{16}$ has a numerator, a fraction-bar, and a denominator.
D. 

There is a 'five' on the slate.

A B C D (c) A. '4' is an even number.
B. '4' is a numeral for 4.
C. 4 is not a number.
D. 4 is a numeral for '4'.

A B C D (d) A. 2 + 2 is the sum of '3' and '1'.
B. '2 + 2' is the sum of 3 and 1.
C. 2 + 2 is a name for '2 + 2'.
D. '2 + 2' is a name for 3 + 1.

2. Let us agree that the number $^{+}2$ measures a 2-miles-to-the-east trip. Then $^{-}2$ measures a 2-mile trip in the opposite direction. Fill in the blanks to make the statements true.



(Example) A trip from D to E is measured by $^{-}4$.

(a) A trip from I to B is measured by _____.

(b) A trip from E to C is measured by _____.

(c) A trip from A to F is measured by _____.

(d) A trip from L to H is measured by _____.

3. **Fill in the blanks to make true sentences.**

(Sample) $^{+}3 + {}^{-}4 = \underline{\;^{-}1\;}$

(a) $^{-}2.7 + {}^{+}8.3 = \underline{\qquad}$

(b) $^{+}4.3 + {}^{+}5.9 = \underline{\qquad}$

(c) $\frac{^{+}3}{5} + \frac{^{-}2}{5} = \underline{\qquad}$

(d) $\frac{^{-}1}{3} + \frac{^{-}1}{2} = \underline{\qquad}$

4. **Fill in the blanks to make true sentences.**

(Sample) $^{+}8 + \underline{\;^{-}8\;} = 0$

(a) $^{-}3.2 + \underline{\qquad} = {}^{+}3.8$

(b) $\underline{\qquad} + \frac{^{+}1}{4} = \frac{^{-}1}{2}$

(c) $^{-}2 + {}^{-}11 = {}^{-}(\underline{\qquad} + 11)$

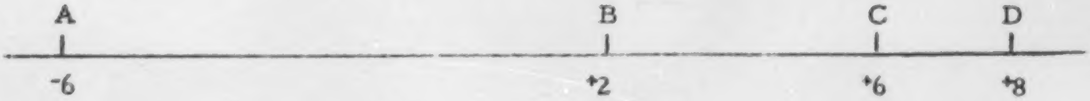
(d) $^{+}8 + \underline{\qquad} = {}^{+}(8 - 3)$

5.

Suppose that a trip on a number line from point A to point B is measured by $^{+}6$ and a trip from point B to point C is measured by $^{-}10$. If the point A is $^{+}7$, what real number is point C? Circle the correct answer.

(A) $^{+}13$ (B) $^{+}4$ (C) $^{+}3$ (D) $^{-}3$ (E) $^{-}4$

6.



A cyclist and a hiker both start moving at the same time and in the positive direction. The cyclist starts at A and the hiker starts at B. On the number line above, which can be used to represent the movements of the cyclist and hiker, point A is located at $^{-}6$, and point B at $^{+}2$. The cyclist passes the hiker at C ($^{+}6$), continues in the positive direction until he reaches a certain point E (not shown) at which time he decides to return to A. As he returns, he passes the hiker at D ($^{+}8$). If both travelers are moving at steady rates, what is the location of the point E at which the cyclist turned? Circle the correct answer.

(A) $^{+}9$ (B) $^{+}10$ (C) $^{+}11$ (D) $^{+}12$ (E) $^{+}13$

7. **Fill in the blanks to make true sentences.**

(Sample) $^{+}5 \times \underline{\;^{+}4\;} = {}^{+}20$

(a) $^{-}3 \times$ _____ $= {}^{+}18$

(b) _____ $\times {}^{-}7 = {}^{-}21$

(c) $^{+}3 \times$ _____ $= {}^{-}1$

(d) _____ $\times {}^{+}3 = 0$

8. **Fill in the blanks to make true sentences.**

(Sample) $^{+}3 \times {}^{+}4 = \underline{\;^{+}12\;}$

(a) $^{-}100 \times {}^{+}\frac{5}{2} =$ _____

(b) $^{-}5 \times {}^{-}\frac{3}{10} =$ _____

(c) $({}^{+}2 \times {}^{-}3) \times {}^{-}4 =$ _____

(d) $^{-}6 \times ({}^{+}2 \times {}^{-}5) =$ _____

9. Complete each line by the expressions given to show the conventional order of performing the operations.

(Example) $3 + 5 \times 2 =$ $3 + (5 \times 2)$

(a) $2 + 8 + 3 =$ ____________

(b) $2 \times 4 - 3 =$ ____________

(c) $6 \times 8 \div 2 =$ ____________

(d) $12 - 5 \times 2 =$ ____________

10. Fill in the blanks to make true sentences.

(Example) $2 \times 5 \times 8 =$ 80

(a) $3 \times 6 - 2 \times 5 + 4 \div 2 =$ ____

(b) $18 - (2 + 6) - 8 \div 2 =$ ____

(c) $7 \times (8 + 4) \div (6 \div 2) =$ ____

(d) $13 - 3 \times (8 \div 4 \times 2) =$ ____

11 Fill in the blanks to make true sentences.

(a) $19 + \underline{\qquad} = 72 + 19$

(b) $(6 + 3) + 2 = 6 + (\underline{\qquad} + 2)$

(c) $2 \times 7 + \underline{\qquad} \times 7 = (2 + 3) \times 7$

(d) $4 \times (5 + 8) = (\underline{\qquad} + 8) \times 4$

12. Here are five principles for real numbers:

(A) Commutative principle for addition (CPA)
(B) Commutative principle for multiplication (CPM)
(C) Associative principle for addition (APA)
(D) Associative principle for multiplication (APM)
(E) Distributive principle for multiplication over addition (DPMA)

Each of the following sentences is an instance of one of the principles listed above. Circle the correct letter to indicate the principle which is used.

A B C D E (a) $(9.837 + 4.652) \times 317.589$
$= (9.837 \times 317.589) + (4.652 \times 317.589)$

A B C D E (b) $(1 + {}^-4) \times (6 \times {}^-17) = (1 + {}^-4) \times 6 \times {}^-17$

A B C D E (c) ${}^+17 + ({}^-12 + {}^-9) = ({}^-12 + {}^-9) + {}^+17$

A B C D E (d) $(\frac{7}{4} - \frac{1}{3}) \times (\frac{1}{2} + \frac{5}{4}) = (\frac{1}{2} + \frac{5}{4}) \times (\frac{7}{4} - \frac{1}{3})$

13. **Fill in the blanks to make true sentences.**

(Sample) ${}^{+}6 - {}^{+}2 =$ ${}^{+}4$

(a) ${}^{+}12 - {}^{+}13 =$ _____

(b) ${}^{-}12 - {}^{+}13 =$ _____

(c) ${}^{+}12 - {}^{-}10 =$ _____

(d) ${}^{-}12 - {}^{-}10 =$ _____

14. **Fill in the blanks to make true sentences.**

(Sample) ${}^{+}6 - {}^{+}2 =$ ${}^{+}4$

(a) ${}^{+}\frac{2}{3} - {}^{+}\frac{1}{6} =$ _____

(b) ${}^{-}\frac{3}{4} - {}^{-}\frac{7}{8} =$ _____

(c) ${}^{-}9.6 - {}^{+}11.3 =$ _____

(d) ${}^{+}17.1 - {}^{-}19.3 =$ _____

15. **The opposite of ${}^{+}4$ is ${}^{-}4$ and the opposite of ${}^{-}4$ is ${}^{+}4$. The operation of oppositing is written by a minus sign "$\underline{-}$". Thus,**

$$-{}^{+}4 = {}^{-}4, \quad -{}^{-}4 = {}^{+}4 \text{ and } --{}^{+}4 = -{}^{-}4 = {}^{+}4$$

Circle 'T' for 'True', 'F' for 'False', or '?' for 'Don't know' or 'Can't tell' for each of the following sentences.

T F ? (a) $---{}^{-}3 = {}^{-}3$

T F ? (b) $-{}^{-}7 + -{}^{+}5 = {}^{-}2$

T F ? (c) $-({}^{-}3 + {}^{-}5) = -{}^{-}3 + -{}^{-}5$

T F ? (d) $-({}^{+}15 \times {}^{-}3) = {}^{-}15 \times {}^{+}3$

16. Circle the letter of the expression which is most appropriate for filling in the blank.

A B C D (a) The sum of positive 7 and negative 3 is ______.
(A) $7 - 3$ (B) ${}^{+}7 - 3$ (C) $7 - {}^{-}3$ (D) ${}^{+}7 + {}^{-}3$

A B C D (b) The difference of negative 5 from negative 3 is ______.
(A) ${}^{-}3 + {}^{-}5$ (B) ${}^{-}3 - {}^{-}5$ (C) $-3 - 5$ (D) $-3 - 5$

A B C D (c) The sum of negative 5 and the opposite of negative 7 is ______.
(A) ${}^{-}5 + -{}^{-}7$ (B) ${}^{-}5 + {}^{-}7$ (C) ${}^{-}5 - {}^{-}7$ (D) $+{}^{-}5 - {}^{-}7$

A B C D (d) The difference of the opposite of negative 4 from 9 is ______.
(A) ${}^{+}9 - {}^{-}4$ (B) $9 - -4$ (C) $9 - +{}^{-}4$ (D) $9 - -{}^{-}4$

17. **Fill in the blanks to make true sentences.**

(Sample) $^{+}15 \div {}^{+}3 =$ $\underline{^{+}5}$

(a) $^{+}8 \div {}^{-}2 =$ _____

(b) $^{-}12 \div {}^{+}4 =$ _____

(c) $^{-}12 \div$ _____ $= {}^{-}2$

(d) _____ $\div {}^{-}5 = {}^{+}4$

18. **Fill in the blanks to make true sentences.**

(Sample) $^{+}6 \div {}^{+}2 =$ $\underline{^{+}3}$

(a) $^{+}1 \div {}^{-}3 =$ _____

(b) $^{+}1 \div \frac{^{-}2}{3} =$ _____

(c) $-\frac{6}{-2} =$ _____

(d) $-\frac{-9}{-3} =$ _____

19. Circle 'T' for 'True', 'F' for 'False', or '?' for 'Don't know' or 'Can't tell' for each of the following sentences. ('<' means 'is less than' and '>' means 'is greater than'.)

(Examples) [✓ T] F ? $^{+}8 < {}^{+}11$
[✓ T] F ? $^{+}15 > {}^{+}10$

T F ? (a) $^{+}2 < {}^{-}3$

T F ? (b) $^{-}2 < {}^{-}3$

T F ? (c) $.01 > .001$

T F ? (d) $\frac{1}{8} > \frac{1}{7}$

20. Which of the following is true for each pair of numbers.

(A) The first member of the pair is greater than the second.
(B) The second member of the pair is greater than the first.
(C) Both members are the same.
(D) We cannot tell whether one is greater than or equal to the other.

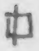
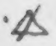
(Example) [✓ A] B C D $(3, 2)$

A B C D (a) $(-\frac{1}{3}, -\frac{1}{7})$

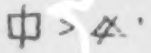
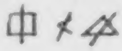
A B C D (b) $(-10, \frac{1}{10})$

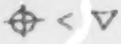
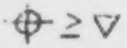
A B C D (c) $(.05, .005)$

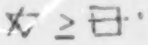
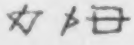
A B C D (d) $(-1, \frac{-3}{3})$

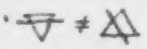
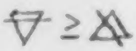
21. The '', '', and other numerals used in the sentences shown below are from the planet Glox. We know that each numeral names a number, but we don't know which number corresponds to any particular numeral. (The signs '>', '$\leq$', etc., have their usual Earth meanings.)

Circle 'T' for 'True', 'F' for 'False', or '?' for 'Don't know' or 'Can't tell' for each of the following sentences.

T F ? (a) If '' is true, '' is _____.

T F ? (b) If '' is true, '' is _____.

T F ? (c) If '' is true, '' is _____.

T F ? (d) If '' is true, '' is _____.

22. The absolute value of ⁺2 is written $|^+2|$, and the absolute value of ⁻2 is written $|^-2|$.

$$|^+2| = 2 \text{ and } |^-2| = 2$$

Fill in the blanks as the sample shows.

(Sample) $|^-3| + |^+7| =$ 3 + 7 = 10

(a) $|^-5| + |^+12| =$ ________________

(b) $|^-5| - |^-12| =$ ________________

(c) $|^-2| \times |^+3| =$ ________________

(d) $|^+18| \div |^-6| =$ ________________

23. In the space at the right of each line, list the two numbers which would correctly complete the statement given.

(Sample) $7 + |\underline{\qquad}| = 10$ $\underline{{}^{+}3,\ {}^{-}3}$

(a) $|{}^{-}9| \times |\underline{\qquad}| = 36$ __________

(b) $|\underline{\qquad}| - |{}^{-}9| = 5$ __________

(c) $|\underline{\qquad} + {}^{-}5| = 10$ __________

(d) $|\underline{\qquad} - {}^{-}5| = 10$ __________

24. In each of the following sentences circle the letter of the expression which is most appropriate for filling the blank. If the answer cannot be determined or is not among the other choices given, circle D.

(A) a positive number
(B) a negative number
(C) 0
(D) cannot tell

A B C D (a) The sum of a negative number and a nonpositive number is _____.

A B C D (b) The sum of a negative number and a positive number is _____.

A B C D (c) The sum of 0 and a nonpositive number is _____.

A B C D (d) Suppose that two numbers are added together, and a third number is added to their sum. If the result is ${}^{+}82$, at least one of the three numbers is _____.

25. Circle 'T' for 'True', 'F' for 'False', or '? for 'Don't know' or 'Can't tell' for each of the following questions.

Suppose that

$$ {}^{+}4 * {}^{+}8 = \frac{1}{2}({}^{+}4 + {}^{+}8) = {}^{+}6, $$

$$ {}^{-}5 * {}^{+}2 = \frac{1}{2}({}^{-}5 + {}^{+}2) = \frac{{}^{-}3}{2}, $$

and

$$ {}^{-}3 * {}^{-}7 = \frac{1}{2}({}^{-}3 + {}^{-}7) = {}^{-}5. $$

Is it also true that:

T F ? (a) ${}^{-}6 * {}^{+}2 = {}^{+}2 * {}^{-}6$

T F ? (b) $({}^{+}1 * {}^{-}4) * {}^{+}2 = {}^{+}1 * ({}^{-}4 * {}^{+}2)$

T F ? (c) ${}^{+}8 * 0 = {}^{+}8$

T F ? (d) ${}^{+}4 * {}^{-}4 = 0$

APPENDIX D

TEST II

(Add 25 to each item number to locate corresponding item data in the preceding tables.)

1. For each of the following open sentences, put a '${}^{+}3$' in each '$\square$' and a '${}^{-}2$' in each '$\bigcirc$'.
If the resulting sentence is true, circle 'T'.
If the resulting sentence is false, circle 'F'.
If the resulting sentence is neither true nor false, circle 'N'.

(Example) [✓ T] F N $\square \times \bigcirc = {}^{-}6$ (because ${}^{+}3 \times {}^{-}2 = {}^{-}6$)

T F N (a) $({}^{+}2 \times \square) - {}^{+}2 = {}^{+}5$

T F N (b) ${}^{+}2 + \bigcirc = ({}^{+}3 \times \square) - {}^{+}9$

T F N (c) ${}^{-}2 \times (\bigcirc + \square) = \bigcirc$

T F N (d) ${}^{+}4 \times ({}^{+}2 + \square) = (\bigcirc \times \square) + (\bigtriangledown + {}^{+}14)$

2. Here is an example of a table of values for the expression '${}^{+}1 - {}^{+}2t$'.

t	${}^{-}1$	${}^{+}2$	0
${}^{+}1 - {}^{+}2t$	${}^{+}3$	${}^{-}3$	${}^{+}1$

Fill in the blanks in the table of values below.

	(a)	(b)	(c)	(d)
m	${}^{+}2$	${}^{+}6$		
p	${}^{-}3$		${}^{-}6$	0
${}^{+}2m + {}^{+}3p$		${}^{+}21$	${}^{-}4$	${}^{+}1$

4. **Complete each sentence to make it true, writing the simplest expression you can in the blank.**

(Example) For each x, the sum of $(x + {}^{+}1)$ and $({}^{+}3x + {}^{+}4)$ is $^{+}4x + {}^{+}5$.

(a) For each x, the product of $^{+}5$ by $^{-}3x$ is __________.

(b) For each x, the difference of $^{+}7$ from $(x + {}^{+}7)$ is __________.

(c) For each x, for each y, $({}^{+}3x + {}^{+}2y)$ exceeds $({}^{+}5x - {}^{+}6y)$ by __________.

(d) For each x, for each y, for each $z \neq 0$, the quotient of $({}^{+}3xz - {}^{+}3yz)$ by $^{+}3z$ is __________.

5. **Complete each sentence to make it true, writing the simplest expression you can in the blank.**

(Example) For each whole number x of arithmetic, if one pencil costs $^{+}2$ cents, x pencils cost $^{+}2x$ cents.

(a) For each whole number x of arithmetic, if one pencil costs $^{+}3$ cents, $(x + {}^{+}5)$ pencils cost __________ cents.

(b) For each number x of arithmetic, if $^{+}3x$ melons are to be distributed equally among $^{+}5$ persons, then each person should receive __________ melons.

(c) For each nonzero number x of arithmetic, a car traveling at a steady rate of $^{+}3x$ miles an hour will travel $^{+}150$ miles in __________ hours.

(d) For each number x of arithmetic, if there are $^{+}75$ sheets of paper in a pile $^{+}1$ inch thick then there are __________ sheets of paper in a pile $^{+}2x$ inches thick.

6. Simplify each of the following expressions:

(Example) ${}^{+}2b + {}^{+}3b =$ ${}^{+}5b$

(a) ${}^{+}5t + {}^{+}8 - {}^{+}2t + {}^{+}8 =$ ______

(b) ${}^{+}5p + ({}^{+}2 - {}^{+}7p) =$ ______

(c) ${}^{+}2({}^{+}7 - {}^{+}3k) + {}^{+}2k =$ ______

(d) $({}^{+}3 - {}^{+}2j) - {}^{+}4({}^{+}2 - j) =$ ______

7. Circle the letter of the expression which, if written in the blank, would give you a true sentence.

(Example) For each y, $y + {}^{+}2y =$ ______
(A) $y \cdot {}^{+}2 \cdot y$ (B) ${}^{+}2 \cdot y \cdot y$ [✓ (C)] ${}^{+}3y$ (D) $y \cdot y \cdot y$

(a) For each x, for each y, ${}^{+}4x \cdot {}^{+}2y =$ ______
(A) ${}^{+}6xy$ (B) ${}^{+}8xy$ (C) ${}^{+}2y + {}^{+}4x$ (D) ${}^{+}8 + xy$

(b) For each x, $({}^{+}3x + {}^{+}7) - ({}^{+}3x - {}^{+}7) =$ ______
(A) ${}^{+}6x$ (B) ${}^{-}6x + {}^{+}14$ (C) 0 (D) ${}^{+}14$

(c) For each y, $(y - {}^{+}1)(y - {}^{+}1) =$ ______
(A) $yy + {}^{+}1$ (B) $yy - y + {}^{+}1$ (C) $yy - {}^{+}2y - {}^{+}1$ (D) $yy - {}^{+}2y + {}^{+}1$

(d) For each x, for each y, $(x - {}^{+}2)(y + {}^{+}3) =$ ______
(A) $xy - {}^{+}6$ (B) $xy - {}^{+}2y + {}^{+}3x - {}^{+}6$ (C) $x + y + {}^{+}1$ (D) ${}^{+}3x - {}^{+}2y$

8. **Simplify each of the following expressions:**

(Example) ${}^{+}(\frac{1}{2}) \div {}^{+}2 =$ ${}^{+}(\frac{1}{4})$

(a) ${}^{+}(\frac{3}{5}) \div {}^{+}(\frac{7}{10}) =$ __________

(b) ${}^{+}0.3 \div {}^{+}0.01 =$ __________

(c) $\frac{{}^{+}4 \cdot {}^{-}2}{{}^{+}3 \cdot {}^{-}4 \cdot {}^{-}6} \div \frac{{}^{+}2 \cdot {}^{-}2}{{}^{+}9 \cdot {}^{+}4} =$ __________

(e) $\dfrac{\dfrac{{}^{+}9}{{}^{+}8 \cdot {}^{-}3 \cdot {}^{-}2}}{\dfrac{{}^{+}3 \cdot {}^{+}5}{{}^{+}4 \cdot {}^{-}2}} =$ __________

9. **Simplify each of the following expressions:**

(Example) ${}^{+}10z \cdot \frac{{}^{+}1}{z} =$ ${}^{+}10$ **provided** $z \neq 0$

(a) $\frac{{}^{+}3x}{{}^{+}2(a + b)} \cdot {}^{+}8x(a + b) =$ __________ **provided** $a + b \neq 0$

(b) $\frac{{}^{+}3xz + {}^{+}6yz}{{}^{+}3z} =$ __________ **provided** $z \neq 0$

(c) $\frac{{}^{+}9x(y + z) - {}^{+}3u(y + z)}{{}^{+}3(y + z)} =$ __________ **provided** $y + z \neq 0$

(d) $\dfrac{\dfrac{{}^{+}1}{{}^{+}5r} + \dfrac{{}^{+}2}{{}^{+}3r}}{\dfrac{{}^{+}1}{{}^{+}2r} - \dfrac{{}^{+}1}{{}^{+}15r}} =$ __________ **provided** $r \neq 0$

10. Each of the statements below is a consequence of one of the following principles for real numbers. Circle the letter corresponding to the principle of which the statement is a consequence.

(A) Commutative principle for addition
(B) Commutative principle for multiplication
(C) Associative principle for addition
(D) Associative principle for multiplication
(E) Distributive principle for multiplication over addition

(Example) [✓ A] B C D E For each x, $x + {}^{+}3 = {}^{+}3 + x$

A B C D E (a) For each x, $({}^{+}3x)(x + {}^{+}5) = (x + {}^{+}5)({}^{+}3x)$

A B C D E (b) For each y, ${}^{+}7 + (y + {}^{+}3) + {}^{-}5 = (y + {}^{+}3) + {}^{+}7 + {}^{-}5$

A B C D E (c) For each x, $({}^{+}2x + {}^{-}1)({}^{+}3x + {}^{+}7)$
$= ({}^{+}2x)({}^{+}3x + {}^{+}7) + ({}^{-}1)({}^{+}3x + {}^{+}7)$

A B C D E (d) For each y, $(y + {}^{+}4)(y + {}^{-}3)(y + {}^{-}1)$
$= (y + {}^{+}4)[(y + {}^{-}3)(y + {}^{-}1)]$

11. Each of the statements below is a consequence of one of the following principles for real numbers. Circle the letter corresponding to the principle of which the statement is a consequence.

(A) Principle for adding 0
(B) Principle for multiplying by 0
(C) Principle for multiplying by ${}^{+}1$
(D) Principle of opposites, or Introduction principle for oppositing
(E) Principle of quotients, or Introduction principle for division
(F) Principle for subtraction, or Definition principle for subtraction

(Example) A B [✓ C] D E F ${}^{-}13 \cdot {}^{+}2 \cdot {}^{+}1 = {}^{-}13 \cdot {}^{+}2$

A B C D E F (a) $({}^{-}3 + 0) + {}^{+}7 = {}^{-}3 + {}^{+}7$

A B C D E F (b) ${}^{+}578 + -{}^{+}578 = 0$

A B C D E F (c) ${}^{+}7 + {}^{-}2 = [({}^{+}7 + {}^{-}2) \div {}^{-}3] \cdot {}^{-}3$

A B C D E F (d) $({}^{-}19 + {}^{+}7) - {}^{+}15 = ({}^{-}19 + {}^{+}7) + -{}^{+}15$

12. Here is a test pattern for a generalization. In each of the first five steps one of the following principles for real numbers is used. The first is given as an example. Circle the correct letter to indicate the principle used in the next four steps.

(A) Commutative principle for addition
(B) Commutative principle for multiplication
(C) Associative principle for addition
(D) Associative principle for multiplication
(E) Distributive principle for multiplication over addition

$y \cdot {}^{+}2 + ({}^{+}5 + {}^{+}3y) + {}^{-}8$
}--- [✓ A] B C D E --- (Example)
$= y \cdot {}^{+}2 + ({}^{+}3y + {}^{+}5) + {}^{-}8$
}--- A B C D E
$= (y \cdot {}^{+}2 + {}^{+}3y) + {}^{+}5 + {}^{-}8$
}--- A B C D E
$= ({}^{+}2y + {}^{+}3y) + {}^{+}5 + {}^{-}8$
}--- A B C D E
$= ({}^{+}2 + {}^{+}3)y + {}^{+}5 + {}^{-}8$
}--- A B C D E
$= {}^{+}5y + ({}^{+}5 + {}^{-}8)$

$= {}^{+}5y + {}^{-}3$

13. Below is a proof of the theorem:

For each x, y, and z, if $z + x = z + y$ then $x = y$.

Fill in the blanks to complete the proof, using one of the following in each blank:

$x,\ y,\ z,\ -x,\ -y,\ -z,\ 0.$

[Proof]

(Example)
Suppose that $z + x =$ __z__ $+ y$ [Given statement]

Then, $z + x + -z = z + y +$ _____

$x + z + -z = y +$ _____ $+ -z$ [Commutative principle for addition]

$x + (z + -z) =$ _____ $+ (z + -z)$ [Associative principle for addition]

$x + 0 = y +$ _____. [Principle of opposites, or Introduction principle for oppositing]

So, $x = y.$ [Principle for adding 0]

Hence, if $z + x = z + y$ then $x = y$.

14. Here are four incomplete lists of some of the ordered pairs of numbers which belong to certain common operations. Complete the lists by inserting the appropriate numerals.

(Example) $(^{+}2, ^{+}4)$, $(^{+}5, ^{+}7)$, $(^{+}11, ^{+}13)$, $(^{+}10, \underline{^{+}12})$

(a) $(^{+}2, ^{+}4)$, $(^{+}3, ^{+}6)$, $(^{+}18, ^{+}36)$, $(^{+}21, ^{+}42)$, $(^{+}10, \_\_\_)$

(b) $(^{+}8, ^{+}5)$, $(^{+}7, ^{+}4)$, $(^{-}3, ^{-}6)$, $(0, ^{-}3)$, $(^{+}3, \_\_\_)$

(c) $(^{+}12, ^{+}4)$, $(^{+}15, 5)$, $(0, 0)$, $(^{-}21, ^{-}7)$, $(\_\_\_, ^{+}6)$

(d) $(^{+}2, ^{+}4)$, $(^{-}4, ^{+}16)$, $(^{+}5, ^{+}25)$, $(^{-}10, ^{+}100)$, $(^{-}3, \_\_\_)$

15. Circle the letter for the ordered pair of numbers which completes the sentence to make it true.

(Example) A B [✓ C] D Addition of 2 contains _____:
(A) (3, 6) (B) (5, 3)
(C) (7, 9) (D) $(\frac{1}{2}, \frac{1}{4})$

A B C D (a) Multiplication by 2 contains _____.
(A) (8, 10) (B) (30, 15) (C) (0, 0) (D) (5, 3)

A B C D (b) Division by 5 contains _____.
(A) (8, 1.6) (B) (10, 5) (C) (12, 2.2) (D) (0, 5)

A B C D (c) The inverse of addition of 6 contains _____.
(A) (6, 1) (B) (7, 13) (C) (50, 46) (D) (70, 64)

A B C D (d) The inverse of multiplication by 7 contains _____.
(A) (14, 21) (B) (0, 0) (C) (7, 0) (D) (5, 35)

16. Pairs of real numbers whose sum is 0 are called opposites, and each is the opposite of the other. Circle the best answer.

(a) What is the opposite of 0? 0 ${}^{+}1$ ${}^{-}1$ none of these

(b) What is the opposite of ${}^{-}1$? 0 ${}^{+}1$ ${}^{-}1$ none of these

(c) What is the opposite of ${}^{+}0.25$? ${}^{+}4$ ${}^{+}0.25$ ${}^{-}0.25$ none of these

(d) What is the opposite of ${}^{-}(\frac{2}{3})$? ${}^{+}(\frac{2}{3})$ ${}^{+}(\frac{3}{2})$ ${}^{-}(\frac{3}{2})$ none of these

17. Pairs of real numbers whose product is ${}^{+}1$ are called reciprocals, and each is the reciprocal of the other. Circle the best answer.

(a) What is the reciprocal of 0? 0 ${}^{+}1$ ${}^{-}1$ none of these

(b) What is the reciprocal of ${}^{-}1$? 0 ${}^{+}1$ ${}^{-}1$ none of these

(c) What is the reciprocal of ${}^{+}0.25$? ${}^{+}4$ ${}^{+}0.25$ ${}^{-}0.25$ none of these

(d) What is the reciprocal of ${}^{-}(\frac{2}{3})$? ${}^{+}(\frac{2}{3})$ ${}^{+}(\frac{3}{2})$ ${}^{-}(\frac{3}{2})$ none of these

18. On each line, circle the letter of the word which would make a true statement if put in the blank. If the correct word is not in the list, circle '(D)'.

(A) inverse
(B) opposite
(C) reciprocal
(D) none of these

A B C D (a) $^-7$ is the _____ of $^+7$.

A B C D (b) $^+0.5$ is the _____ of $^+2$.

A B C D (c) $\frac{^+1}{^+3}$ is the _____ of $^-3$.

A B C D (d) Subtraction of $^+4$ is the _____ of addition of $^+4$.

19. On each line, circle the letter of the number which would make a true statement if a numeral for it were put in the blank. If the correct number is not in the list, circle '(D)'.

(A) $^+1$
(B) 0
(C) $^-1$
(D) none of these

A B C D (a) The result of dividing any number by 0 is _____.

A B C D (b) The result of multiplying any number by _____ is the opposite of the number.

A B C D (c) The reciprocal of _____ is 0.

A B C D (d) The opposite of the reciprocal of _____ is $^-1$.

20. Circle one correct answer for each question.

(a) How many real numbers satisfy the sentence '$a + {}^{+}5 = {}^{+}2$' ?

(A) none (B) only 1 (C) only 2 (D) all

(b) How many real numbers satisfy the sentence '$aa = 0$' ?

(A) none (B) only 1 (C) only 2 (D) all

(c) How many real numbers satisfy the sentence '$bb = {}^{-}1$' ?

(A) none (B) only 1 (C) only 2 (D) all

(d) How many real numbers satisfy the sentence '$(p - 3)(p + 2) = 0$' ?

(A) none (B) only 1 (C) only 2 (D) all

21. For each of the following sentences, circle 'T' for 'True', 'F' for 'False', and '?' for 'Can't tell'.

T F ? (a) Some open sentences are true.

T F ? (b) A false sentence is not a statement.

T F ? (c) The two expressions '${}^{+}5a + {}^{-}2a$' and '${}^{+}3a$' are equivalent expressions.

T F ? (d) A counter-example to a universal generalization shows that the generalization is false.

22. In order to write generalizations concisely, we use the symbol '$\forall$'. Write each of the generalizations below using '$\forall$' and other algebraic symbols.

(Sample) No matter what real number you pick, its product by 0 is 0.

$$\forall_x\ x \cdot 0 = 0$$

(a) Whatever real number you choose, if you subtract the number from 0, you get the opposite of the number.

(b) For each real number, the result of adding its product by $^{+}7$ to its product by $^{-}3$ is the same as its product by $^{+}4$.

(Sample) No matter what nonzero real number you pick, the quotient of the number by itself is $^{+}1$.

$$\forall_{x \neq 0}\ \frac{x}{x} = {}^{+}1$$

(c) Pick any real number different from $^{-}1$. Add $^{+}1$ to it. Double the result and then divide by the sum of the number you picked and $^{+}1$. The final result is $^{+}2$.

(d) Pick any real number x. Now pick a second real number y which is not the opposite of x. It will always turn out that

$$\frac{x}{x+y} - \frac{y}{x+y} = \frac{x-y}{x+y}$$

23. The open sentence '$\square + {}^{+}2 = {}^{+}4$' becomes a true statement if we put a '${}^{+}2$' in place of the '$\square$'. In other words, it is true that there exists a real number such that the sum of it and ${}^{+}2$ is ${}^{+}4$. For simplicity, we use the symbol '$\exists$' to express the sentence above, and write:

$$\exists_x \ (x + {}^{+}2) = {}^{+}4$$

Translate each of the following sentences into a sentence which uses a '$\exists$' and other algebraic symbols.

(a) There is a real number such that the product of that number by ${}^{-}7$ is ${}^{+}35$.

____

(b) There exists a real number whose sum with itself is 0.

____

Furthermore, we can use both of the symbols '$\forall$' and '$\exists$' to write a concise statement of the generalization that, for each first real number, there is a second real number such that the product of the first by itself is the second. Here is such a concise statement:

$$\forall_x \ \exists_y \ x \cdot x = y$$

Translate each of the following sentences into a sentence which uses '$\forall$', '$\exists$', and other algebraic symbols.

(c) For each first real number, there is a second real number such that the sum of ${}^{-}7$ and the second is the first.

____

(d) There exists a first real number such that, for each second real number, the sum of ${}^{+}6$ and the second is the first.

____

24. On each line, circle the letter of the answer which would make a true statement if put in the blank. If answers A, B, and C would all make true statements, circle 'D'. If neither the A, B, or C answers would make true statements, circle 'E'.

(A) $\forall_x$ if x is positive then
(B) $\forall_x$ if x is negative then
(C) $\forall_x$ if $x \neq 0$
(D) All of these
(E) None of these

(Example) A B C [✓ D] E ____________ $(x + {}^{+}1) - x$ is positive.

A B C D E (a) ____________ $-x$ is negative.

A B C D E (b) ____________ $x \cdot x$ is positive.

A B C D E (c) ____________ $x \cdot -x$ is nonpositive.

A B C D E (d) ____________ $(x \cdot x) - (x + {}^{+}1)$ is nonnegative.

25. On each line circle the letter of the answer which would make the most general true statement possible.

(Example) A B C D __________ $x + y = y + x$

[✓ (A)] $\forall_x \forall_y$ (B) $\forall_{x \neq 0} \forall_y$ (C) $\forall_x \forall_{y \neq 0}$ (D) $\forall_{x \neq 0} \forall_{y \neq 0}$

A B C D (a) __________ $(x - y)(x + y) = xx - yy$

(A) $\forall_x \forall_y$ (B) $\forall_{x \neq 0} \forall_y$ (C) $\forall_x \forall_{y \neq 0}$ (D) $\forall_{x \neq 0} \forall_{y \neq 0}$

A B C D (b) __________ $-z(x + y) + -z(x - y) = -{}^{+}2xz$

(A) $\forall_x \forall_y \forall_z$ (B) $\forall_x \forall_y \forall_{z \neq 0}$ (C) $\forall_{x \neq 0} \forall_{y \neq 0} \forall_z$ (D) $\forall_{x \neq 0} \forall_y \forall_{z \neq 0}$

A B C D (c) __________ $\frac{z}{x} + y = \frac{z + xy}{x}$

(A) $\forall_x \forall_y \forall_z$ (B) $\forall_x \forall_{y \neq 0} \forall_z$ (C) $\forall_{x \neq 0} \forall_y \forall_z$ (D) $\forall_{x \neq 0} \forall_y \forall_{z \neq 0}$

A B C D (d) __________ $\frac{1}{x} - \frac{1}{y} = \frac{y - x}{xy}$

(A) $\forall_x \forall_y$ (B) $\forall_{x \neq 0} \forall_y$ (C) $\forall_x \forall_{y \neq 0}$ (D) $\forall_{x \neq 0} \forall_{y \neq 0}$

26. For each sentence, find a real number which "satisfies" it.

(Example) $^{+}8 + m = {}^{+}12$ $\underline{\ {}^{+}4\ }$

(a) $^{+}3.5 + a = {}^{+}2.5$ __________

(b) $^{+}5 = {}^{+}6b - {}^{+}7$ __________

(c) $^{+}2c + {}^{+}5 = {}^{+}15$ __________

(d) $\frac{{}^{+}5 + d}{{}^{+}4} = {}^{+}6$ __________

27. For each sentence, find all the real numbers which "satisfy" it. If there are none, write 'none'.

(Example) $a \cdot a = {}^{+}16$ $\underline{\ {}^{+}4,\ {}^{-}4\ }$

(a) $b(b - {}^{+}2) = 0$ __________

(b) $c \cdot c + {}^{+}4 = 0$ __________

(c) $d \cdot d \div {}^{+}100 = 0$ __________

(d) $\frac{{}^{-}2e}{e} = 0$ __________

28. Below is a proof of the theorem:

$$\forall_x \; x0 = 0$$

Fill in the blanks to complete the proof, using one of the following in each blank:

$x,\ xx,\ x0,\ 0,\ -(xx)$

(Proof)

[Principle for adding 0]

$x + 0 = x$

[Uniqueness principle for multiplication]

$x(x + 0) =$ __________

[Commutative principle for multiplication]

$(x + 0)x = xx$

[Distributive principle for multiplication over addition]

$xx + 0x = xx$

[Commutative principle for multiplication]

$xx + x0 = xx$

[Principle for adding 0]

$xx + x0 = xx +$ __________

[Uniqueness principle for addition]

$xx + x0 +$ __________ $= xx + 0 + -(xx)$

[Commutative principle for addition]

$x0 + xx + -(xx) = 0 + xx + -(xx)$

[Associative principle for addition]

$x0 + [xx + -(xx)] = 0 + [xx + -(xx)]$

[Principle of opposites or Introduction Principle for oppositing]

$x0 +$ __________ $= 0 + 0$

[Principle for adding 0]

$x0 = 0$

29. Derive the following generalization:

$$\forall_x \forall_y \forall_z \ (x + y) + z = (z + y) + x.$$

In your derivation use only the principles listed below:

Commutative principle for addition [CPA]
Commutative principle for multiplication [CPM]
Associative principle for addition [APA]
Associative principle for multiplication [APM]
Distributive principle for multiplication over addition [DPMA]